Fractional Brownian motion and weather derivatives

by

Qingsheng Dong

THESIS

for the degree of

MASTER OF SCIENCE

(Master i Modellering og dataanalyse)

Faculty of Mathematics and Natural Sciences
University of Oslo

May 2008

Faculty of Mathematics and Natural Sciences
University of Oslo

Contents

List of Figures

List of Tables

Abstract

The fractional Brownian motion (fBm) has recently drawn a lot of attention and has been studied in several directions, such as stochastic integration, stochastic differential equations, financial applications, and solutions for many other theoretical problems. This Master thesis focuses on investigating the financial applications which is built on the fBm platform, and it studies weather derivatives as a classical example. In the first part of this thesis, Wick Itô Skorohod (WIS) integrals are introduced as the stochastic integrals of the financial model based on fBm. To establish a parallel fractional financial model to the well-known Black-scholes model, which is driven by the classical Brownian motion, a fractional version of Itô formula and the Girsanov theorem are presented. The solution of the fractional Ornstein-Uhlenbeck equation is also given in this part. In the second part of this thesis, the weather market is studied in two aspects: on one side, the stochastic model for temperature-based derivatives and its analytical solutions for pricing; and on the other side, data analysis from five Norwegian districts and the Monte Carlo pricing. This thesis tries to give an overall understanding of fBm from the theoretical interest to financial model and real-world significance

Acknowledgment

I have spent half a year on this thesis. That has been an instructive period. I'm grateful for all the people who have had the interesting discussions with me on the topics related to the fractional Brownian motion and the weather derivatives market. Particularly I would like to thank my supervisor, Tom Louis Lindstrøm, for introducing the interesting subjects to me. He has patiently guided me through the mathematical challenges. And also a special thank to Professor Fred Espen Benth for many inspiring discussions. For providing data for me, a thank to Jostein Amdal and Kevan Cheyne, from Gjensidige.

In addition, I would like to thank Haifeng Sun, for the assistance of the language in the thesis. Last but not the least, a warm thank to my family and friends, who have been supporting me through the Master program. I have indeed learned a lot, and will bring all that I've learned into my life and career in the future.

Oslo, May 2009 Qingsheng Dong

Chapter 1

Introduction

The classical Brownian motion (Bm) has been the most successful theoretical model for many applications. The studies of Bm have actually gained impressive results in all major scientific fields such as mathematics, physics, chemistry and biology. However, the sufficiency of the key assumption of Bm which is the independence of increments have been questioned recently. Observations of many long memory phenomena, which cannot be described with Bm, have inspired us to a more generalized class of continuous time Gaussian processes. The fractional Brownian motion (fBm) with a H coefficient $\in (0, 1)$ satisfies the desire to quantify the correlations between increments, in order to model the targeted long memory phenomena more accurately. The H coefficient named after the British hydrologist Harold Edwin Hurst (1880-1978), is the most widely used measurement for the long-range dependence of increments.

The studies of fBm have moved into several directions, just to mention, the two major aspects: financial applications and modeling of nature phenomena such as temperature, solar activity and water level. In this Master thesis, these two aspects are connected via introduction of weather derivatives, which is a new class of financial instruments. The purpose of the weather derivatives market is to provide an alternative strategy to manage the unpredictable weather risk. A brief introduction of weather derivatives market is given together with a real-life example of the car insurance. The data are provided by Gjensidige, which is one of the leading insurance groups in the Nordic general insurance market. The considered weather factor in the this thesis is temperature. Temperature is one of the most significant weather factors for economical activities and this is the reason why temperature-based derivatives are most traded in the market. The studies on temperature dynamic for the purpose of pricing have drawn a lot of attention. The physical reality motivated a stochastic model driven by a long memory, or in other words, long-range dependent process, such as fBm. The fractional Ornstein-Uhlenbeck process is proposed by Brody, Syroka and Zervos [8] and price formulas are derived for the most common temperature indexes, such as HDD, CDD and CAT, by using partial differential equations. Later on, Benth [1] presented an arbitrage-free model for derivatives on temperature, using the same fractional Ornstein-Uhlenbeck process, together with quasi-conditional expectation and Wick Itô Skorohod(WIS) integrals of fBm.

The WIS integrals of fBm are developed by Hu, Øksendal et al.[4]. Dynamics of the option values, are derived inn addition to the prices for contracts in [1]. The mentioned articles or book are the theoretical background of this thesis. This thesis attempt to give a overview of the most important results of stochastic calculus for fBm, such as definition of the WIS integral, fractional Itô formula and fractional Girsanov theorem, as well as to make a sketch of a arbitrage-free model for temperature based derivatives. Since the Black-scholes model is well establish, the market model driving by fBm is introduced as a extension of the Black-scholes model. Many mathematical fields are involved in the process of develop a framework for fBm. For the approach employed in this thesis, the knowledge about stochastic analysis, fractional white noise, Fourier transform, Wiener-Itô expansion and Wick product are vital. The semimartingale issue is the major barrier for a fractional market model, and the barrier is tried to be removed by the introduction of quasi-conditional expectation and quasi-martingale. Analytical price formulas are given in the end of this part.

In the nest part of the thesis, a data analysis is performed on daily temperature of five Norwegian districts, in a period from 1990 to 2008. The purpose is to justify the theoretical dynamic of temperature and study the fractional property of the temperature data. The method of analysis is inspired by Benth and Saltyte-Benth [2]. The discrete time AR(1) model with fractional residuals is suggested as a discrete version of the fractional Ornstein-Uhlenbeck process, and parameters of the model is estimated for the five districts. The parameters involved in AR(1) are trend parameter, seasonality parameter, mean-reverting rate and value of H coefficient. In this part of thesis, a discussion is given on statistic properties of the Norwegian daily temperature. The central issues is normality and the fractional property of the temperature data during the four stages of analysis: original (OR), detrended and deseasonalized (DD), residuals after regression (RES) and residuals divide by seasonal sigma (RES/SIGMA). In order to estimate H values of data, a generator for fBm, based on Wood-Chan's method is introduced, together with three popular estimators for H values: the ST method, the RS method and the DFA method. A Monte Carlo simulation is performed to test the efficiency of the three estimators. Since AR(1) model is a competitor for fBm to capture the long-rang dependence, a higher order AR model is tested. The higher order AR model has not improved the AR(1) model, in the respect of removing the fractional property from residuals.

The last part of the thesis using a Monte Carlo approach to price the temperature-based weather derivatives. Beside the HDD, CDD and CAT indexes, a over-the-counter contract based on number of icing days, is priced for different values of H. The extension to a temperature dynamic driving by fBm, does have significant effect on some types of contracts. However when the selected weather station is Oslo, prices of the most traded HDD, CDD and CAT contracts are not effected by the variation of H values. For the contract types, where the whole temperature evaluation in the contract period are counted, the H values can influence price. The degree of the effect is depended on H values, but also on the level of strike.

The thesis focus on breadth of fBm and it's application in the weather derivatives market. Intension is to give a total understanding of the fBm and the weather market, from the theoretical interest to financial model and real-world significance. The thesis is organized as follow:

- Figures are collectd in Apendix A and R scripts in Apendix B. All R scripts are written as functions and therefore easy to applied. All R scripts begin with a description of the function and a guide of use.

- In chapter 1, the fBm is defined and the WIS integrals of fBm introduced in eight steps. The fractional Itô formula and Girsanov theorem are presented. Using the fractional Itô formula, the fractional Ornstein-Uhlenbeck equation is solved.

- In chapter 2, a brief of the weather market is given with it's important elements. Such as the underlying indexes HDD, CDD and CAT. A real life example from the car insurance, demonstrated the potential of the weather market. At the end of the chapter, the four major methods of pricing are summarized.

- In the chapter 3, the classical Black-scholes model is extended to a fractional version. A arbitrage-free market model for the temperature-based derivatives is established. The analytical solutions for HDD, CDD and CAT under the risk-neutral probability measure($\mathcal{Q}$) are derived.

- In the chapter 4, a data analysis is carried out on the daily temperature of the five Norwegian districts. The analysis in this chapter included parameter estimation, normality analysis and H values estimation. An fBm generator, together with three estimator for H values are used and compared. At the end of this chapter, AR model are compared to model driving by fBm, in the matter of capture long-range dependence.

- In the chapter 5, a Monte Carlo approach is used to price HDD, CDD and CAT. Influence of H values on prices is studied.

Chapter 2

Fractional Brownian motion

The very first article about fractional Brownian motion (fBm) was published in 1940, by Andrey Nikolaevich Kolmogorov (1903-1987), a Soviet Russian mathematician. He introduced continuous time Gaussian processes with stationary increments and with the self-similarity property. Kolmogorov named such processes as 'Wiener Spirals'. However, that was Benoît B. Mandelbrot (1924-), a French mathematician and also best known as the father of fractal geometry. He considered an integral representation for fBm via a classical Brownian motion (Bm), and named the process as 'fractional Brownian motion'. The fBm became a hot topic again in the 1990s. Long-range dependent process received increasing interest in this period due to the insufficiency of the classical driving process with independent increments. Except the case $H = \frac{1}{2}$), the fBm is neither semi-martingale nor a Markov process. The stochastic calculus foundation for Black-scholes model therefore felt apart. A lot of attempts on a generalization of the classical stochastic analysis give remarkable results, mainly in direction of a fractional finance market model. The approach employed in this thesis is based on Wick Itô Skorohod(WIS) integrals, which leads to a arbitrage-free marked model. The focus of the chapter is give a overview of the most important resultants in stochastic calculus of fBm, and organize them in a way simply for further applications. The details of proofs are excluded and can be found in references.

In this chapter, definition of the fBm, together with WIS integral are introduced. Fractional Ornstein-Uhlenbeck equation is solved using fractional Itô formula. Solution of the equation is used to describe the temperature dynamic later in this thesis. The main reference of this chapter are Biagini, Hu and Øksendal et al.[4] and Mishura [19]. This two books attempt to give a systematic overview of existing results for stochastic calculus of fBm. For further reading, the following books or articles are recommended: Biagini, Øksendal, Sulem and Wallner[5] for an introduction to white noise theory, Tomas and Hult[6] for Wick products, Henrik and Øksendal[14] for fractional Brownian motion in Finance and Lourie[16] for Ornstein-Uhlenbeck equations.

2.1 Properties of the fractional Brownian motion

2.1.1 Definition

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space, definition of fBm is given by:

Definition 2.1 (Fractional Brownian motion) *A fractional Brownian motion(fBm)* $B_t^H(t \geq 0)$ *of Hurst coefficient H,* $H \in (0,1)$ *is a continuous and centered Gaussian process with covariance function:*

$$E\left[B_t^H B_s^H\right] = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H})$$

Author refer to [4] for this definition. Note that for $H = \frac{1}{2}$, fBm is a classical Brownian motion(Bm). The fractional Brownian motion has following properties by *Definition 2.1*:

i). $B_t^H = 0$, and $E\left[B_t^H\right] = 0$ for all $t \geq 0$

ii). B_t^H has homogeneous increments. $B_{t+s}^H - B_s^H$ follows the same probability law of B_t^H for $s, t \geq 0$.

iii). For all $s, t \geq 0$, $E(|B_t^H - B_s^H|^2) = |t-s|^{2H}$

The increments of B_t^H also called as fractional Gaussian noise(fGn). In the other words, the fBm is the integral or the cumulative sum of the fGn. B_t^H is a Gaussian process with continuous modifications, this property is guaranteed by Komogorov theorem. Later in this chapter, construction of fBm through the white noise theory will be presented. The most important elementary property of the fBm is the long-range dependence.

2.1.2 Correlation, long-range dependence and other properties

Increments of the classical Brownian motion are independent. Mathematical speaking, correlation function of Bm is zero. Brownian motion is widely used as driving process i stochastic modeling. The popularity of Brownian motion is partly mathematical convenience and partly because of the representativeness of Bm, determined by it's normal distributed, independent increments. Motivation for a extension from Bm to fBm is the same as introduction of the stochastic modeling, which is a desire to explain nature phenomena more precisely. Long-range dependence are observed in finance, teletraffic, and in many nature phenomena, such as water level, solar activity and daily temperature. Specially daily temperature dynamic, driving by fBm, appeared as subject in many studies. The temperature dynamic will be discussed in details in *chapter 4* and *chapter 5*.

The correlation between two increments of a fBm, is determined by Hurst coefficient. For $H \in (0,1)$ and $s < t, s+h < t$, the autocovariance function of fBm, follow by the *properties iii)* is

$$E((B_{t+h}^H - B_t^H)(B_{s+h}^H - B_s^H)) = \frac{1}{2}(|t-s-h|^{2H} + |t-s+h|^{2H} - 2|t-s|^{2H})$$

or in the integral form

$$E((B^H_{t+h} - B^H_t)(B^H_{s+h} - B^H_s)) = 2(H - \frac{1}{2})H \int_s^{s+h} \int_t^{t+h} (u - v)^{2H-2} du dv$$

From the integral form, obviously, increments have positive correlation when $H \in (\frac{1}{2}, 1)$, and negative correlation when $H \in (0, \frac{1}{2})$. Processes in the first case, are persistence. The behavior of a persistent process is aggregated. Changes in the past have a positive influence on present time and in the future. The process have a memory effect. The second case, it is called antipersistence and behavior in the opposite way. The nest property of fBm is long-rang dependence, which is determined by correlation of increments. The mathematical definition of the long-rang dependence is:

Definition 2.2 (Long-range dependence) *A stationary process X_t exhibits long-range dependence if the autocovariance function $R_H(n) := cov(X_k, X_{k+n})$ satisfy*

$$\lim_{n \to \infty} \frac{R_H(n)}{cn^{-\alpha}} = 1$$

for some constant c and $\alpha \in (0, 1)$

Recall the covariance function of fBm. It is defined as follow:

$$R_H(t, s) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}), \;\; s, t \geq 0 \tag{2.1}$$

Use Taylor expansion on $\text{cov}(B^H_s - B^H_{s-1}, B^H_{s+n} - B^H_{s+n-1})$ gives:

$$R_H(n) = \frac{1}{2}\left[(n+1)^{2H} + (n-1)^{2H} - 2n^2H\right] \sim H(2H-1)n^{2H-2}, \;\; |n| \to \infty$$

And therefor

- For $H \in (0, \frac{1}{2}), \sum_{n=1}^{\infty} |R_H(n)| < \infty$
- For $H \in (\frac{1}{2}, 1), \sum_{n=1}^{\infty} R_H(n) = \infty$

The fBm have a long-range dependence property when $H \in (\frac{1}{2}, 1)$, since

$$\lim_{n \to \infty} \frac{R_H(n)}{H(2H-1)n^{2H-2}} = 1$$

An other property of fBm is self-similarity. For $H \in (0, 1)$ and $\alpha > 0$, the law of $B^H_{\alpha t}$ is the same as the law of $\alpha^H B^H_t$. This property is determined by the covariance function as well. The covariance function is homogeneous of order 2H and therefor the fBm is self-similar with order H.

In this thesis, the last property of fBm which should be noticed is, when $H \neq \frac{1}{2}$, the fBm is neither semimartingale nor Markov. The semimartingales form the largest class of processes for which the Itô integral can be defined. In consequence, definition of the stochastic integral for the fBm, need a new approach. In the next section, one of the approaches is introduced.

2.2 Stochastic calculus

As all know, from a deterministic model to a stochastic model, a random 'noise' term is added. The stochastic model is generally of the form.

$$\frac{dX}{dt} = \mu(t, X_t) + \sigma(t, X_t) \cdot W_t \tag{2.2}$$

where W_t is known as 'white noise'. The classical Brownian motion is a typical white noise process. In order to extend the classical Bm model (2.2) to a model driven by fBm, the 'white noise' should be replaced by increments of B_t^H. In other words, the equation need to be solved:

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s^H \tag{2.3}$$

As same as the classical Brownian motion case, a definition is needed for

$$\int_0^t f(s, \omega)dB_s^H(\omega) \tag{2.4}$$

There are several ways to reach the goal. From the view of simulation, the pathwise integration is which makes most sense, but unfortunately a market defined in this way has arbitrage opportunity. As mentioned before, fBm is not semimartingale when $H \neq \frac{1}{2}$. This is the mathematical reason for free lunch with vanishing risk. This is the major problem for a fractional finance market model and the reason why many studies did not suggest fBm as driving process for a market model. The next try is Wiener integrals defined for integrand f as deterministic functions and can be extend to $f(s, \omega)$ as a stochastic process by Skorohod integral. This approach is well defined for the case $H \in (\frac{1}{2}, 1)$, but the pathes of fBm become too irregular to define for $H \in (0, \frac{1}{4})$. The third approach is fractional Wick Itô Skorohod(fWIS) integrals and Wick Itô Skorohod(WIS). WIS and fWIS are developed based on white noise theory. WIS is defined for $H \in (0, 1)$ and there are many useful fractional calculus already been proofed for WIS. The most important of them are fractional Itô isometry and fractional Itô formula. Market defined by WIS is free from arbitrage. Even though WIS is less intuitive by simulation, we choose to use this definition in the theoretical part of the thesis. The biggest consideration by use WIS to defined finance market is that portfolio and price defined this way has no natural economic interpretation. But since weather derivatives market is not complete and we can't buy and hold a temperature measurement any way, WIS is a better choice.

More detail for relations between different integrals w.r.t fBm, can be found in [4]. And [6] gives a comparison between finance markets modeled by pathwise and WIS integrals.

2.2.1 Wick Itô Skorohod(WIS) integrals for fractional Brownian motion

Let $\mathcal{S}(\mathbb{R})$ denote the Schwartz space of rapidly decreasing smooth functions on $\mathbb{R}$, and the WIS integral is defined on probability space $\Omega := \mathcal{S}'(\mathbb{R})$, which is dual of $\mathcal{S}(\mathbb{R})$. $\Omega := \mathcal{S}'(\mathbb{R})$ is the space of tempered distributions. In the rest of the theoretical part, the Schwartz space is employed.

In order to definer WIS, a lot of new definitions and theorems are applied. It's easier if the whole picture is presented before getting into the details. The defining process is divided into 8 steps. Step 1 is just notation. M operator is introduced first in Step 2. Then in Step 3, B_t and B_t^H are defined based on indicator function $I_{[0,t]}$ and M operator. In Step 4, integrals are defined for deterministic integrand. A relation between integrals w.r.t B_t and B_t^H is proofed in step 5. The Wiener-Itô chaos expansion is presented in Step 6. In Step 7, white noise and fractional white noise are introduced. Finally in Step 8, a definition of WIS integrals is given. The results present in this thesis are already developed by R.J. Elliott, Francesca Biagini, Bernt Øksendal and many other mathematicians. The author refer to [5] for details of proofs. This thesis gives just a sketch of construction of WIS, and make the approach simply for understanding. The following Table 2.1 attempt to give a overview.

Table 2.1: *WIS in 8 steps, for Bm and fBm*

	Classical Brownian motion	Fractional Brownian motion
1. Notation	B_t	B_t^H
2. Indicator/Operator	$I_{[0,t]}(s) = \begin{cases} 1 & if\ \ 0 \le s \le t \\ -1 & if\ \ t \le s \le 0, \\ & except\ \ t = s = 0 \\ 0 & otherwise \end{cases}$	$Mf(x) = \begin{cases} C_H \int_{\mathbb{R}} \frac{f(x-t)-f(x)}{\lvert t\rvert^{3/2-H}} dt & 0 < H < \frac{1}{2} \\ f(x) & H = \frac{1}{2} \\ C_H \int_{\mathbb{R}} \frac{f(t)}{\lvert t-x\rvert^{3/2-H}} dt & \frac{1}{2} < H < 1 \end{cases}$ $C_H = \left\{2\Gamma(H-\frac{1}{2})\cos[\frac{\pi}{2}(H-\frac{1}{2})]\right\}^{-1} [\Gamma(2H+1)\sin(\pi H)]^{\frac{1}{2}}$
3. Definition	$B_t =< \omega, I_{[0,t]} >$	$B_t^H =< \omega, M_{[0,t]} >$
4. Integral	$< \omega, f >= \int_{\mathbb{R}} f(t)dB_t$	$< \omega, Mf >= \int_{\mathbb{R}} f(t)dB_t^H$
5. Relation	$\int_{\mathbb{R}} f(t)dB_t^H = \int_{\mathbb{R}} Mf(t)dB_t$	
6. The Wiener-Itô chaos expansion	$B_t = \sum_{k=1}^{\infty} \int_0^t \xi_k(s)ds\mathcal{H}_{\epsilon^{(k)}}(\omega)$	$B_t^H = \sum_{k=1}^{\infty} \int_0^t M\xi_k(s)ds\mathcal{H}_{\epsilon^{(k)}}(\omega)$
7. White noise	$W_t = \sum_{k=1}^{\infty} \xi_k(t)\mathcal{H}_{\epsilon^{(k)}}(\omega)$	$W_t^H = \sum_{k=1}^{\infty} M\xi_k(t)\mathcal{H}_{\epsilon^{(k)}}(\omega)$
8. WIS	$\int_{\mathbb{R}} f(t,\omega)\delta B_t = \int_{\mathbb{R}} f(t,\omega) \diamond W_t dt$	$\int_{\mathbb{R}} f(t,\omega)dB_t^H = \int_{\mathbb{R}} f(t,\omega) \diamond W_t^H dt$

Step 2

The whole idea begin with find a relation between the classical Brownian motion and the fBm. Since calculus for the classical Brownian motion is fully developed and well known, it will be much more intuitive to understand fBm if a operator can be found, which 'turns' a fBm into Bm. The famous M operator does the job.

Definition 2.3 (The M operator) *Let $0 < H < 1$. The operator $M = M_H$*

is defined on functions $f \in \mathcal{S}(\mathbb{R})$ by

$$\widehat{Mf}(y) = |y|^{1/2-H}\hat{f}(y), \ \ y \in \mathbb{R} \tag{2.5}$$

where

$$\hat{g} := \int_{\mathbb{R}} e^{-ixy} g(x)dx$$

denotes the Fourier transform and $\mathcal{S}(\mathbb{R})$ denote the Schwartz space ofb rapidly decreasing smooth functions on $\mathbb{R}$.

The M operator have following properties:

- $Mf(x) = f(x)$ for $H = \frac{1}{2}$
- $MI_{[0,t](x)} := M[0,t](x)$
- $\langle f, Mg\rangle_{L^2(\mathbb{R})} = \langle Mf, g\rangle_{L^2(\mathbb{R})}$
- $M_H(M_{1-H}f) = f, \ f \in \mathcal{S}(\mathbb{R})$
-
$$\int_{\mathbb{R}} M_{[0,t]}M_{[0,s]}(x)dx = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}) \tag{2.6}$$

Step 3

The fBm can be defined by M operator. The approach is similar to definition of Bm by white noise theory. Bm is defined as followed:

$$B_t := B_t(t,\omega) :=< \omega, I_{[0,t](\cdot)} > \tag{2.7}$$

For $t \in \mathbb{R}$

$$\tilde{B}_t^H := \tilde{B}^H(t,\omega) :=< \omega, M_{[0,t]}(\cdot) > \tag{2.8}$$

where $< \omega, f >= \omega(f)$ is the action of $\omega \in \Omega$.
We need to calculate:

- $E(\tilde{B}_t^H)$
- $E(\tilde{B}_t^H \tilde{B}_s^H)$

Definition 2.4 *$\mathcal{S}(\mathbb{R})$ is the Schwartz space of rapidly decreasing smooth functions on $\mathbb{R}$, and let $\Omega := \mathcal{S}'(\mathbb{R})$ be its dual, the space of tempered distributions. Let μ be probability measure on the Borel set $\mathcal{B}(\mathcal{S}'(\mathbb{R}))$ defined by:*

$$\int_{\mathcal{S}'(\mathbb{R})} exp(i < \omega, f >)d\mu = exp(-\frac{1}{2}||f||^2_{L^2(\mathbb{R})}), \ \ f \in \mathcal{S}(\mathbb{R})$$

By *definition 2.4*

$$E[< \omega, f >] = 0$$

and therefor

$$E(\tilde{B}_t^H) = 0, \ \ for \ \Omega = \mathcal{S}'(\mathbb{R})$$

And the expectation $E(\tilde{B}_t^H \tilde{B}_s^H)$ is gives by (2.6)

$$E(\tilde{B}_t^H \tilde{B}_s^H) = \int_{\mathbb{R}} M_{[0,s]}(x) M_{[0,t]}(x) dx = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$$

Therefor the continuous version B_t^H of $\tilde{B}_t^H$ is a fBm.

Step 4

The attempt is now to define the integral (2.4).
Let f be a step function of the form

$$f(t) = \sum_j a_j I_{[t_j, t_{j+1}]}(t)$$

Then

$$\langle \omega, Mf \rangle = \sum_j a_j (B_{t_j+1}^H - B_{t_j}^H) = \int_{\mathbb{R}} f(t) dB_t^H$$

In the other hand

$$\langle \omega, f \rangle = \sum_j a_j (B_{t_j+1} - B_{t_j}) = \int_{\mathbb{R}} f(t) dB_t$$

Step 5

The desired relation is direct result of step 4, it is give as follow:

$$\int_{\mathbb{R}} f(t) dB_t^H = \int_{\mathbb{R}} Mf(t) dB_t \tag{2.9}$$

The relation (2.9) indicates a way to 'uncorrelated' B_t^H into B_t, and using a similar approach, a extension from Itô integral and it's related theorems to a fractional version is realistic.

Step 6

Definition 2.5 (The Wiener-Itô chaos expansion theorem I) *La $F \in L^2(\mu)$. Then there exists a unique sequence $\{f_n\}_{n=0}^{\infty}$ of functions $f_n \in L^2(\mathbb{R}^n)$, such that*

$$F(t) = \sum_{n=0}^{\infty} I_n(f_n), \quad \textit{where } f_n \textit{ are symmetric deterministic}$$

where

$$\begin{aligned} I_n(f_n) &:= n! \int_{\mathbb{R}} \cdots \int_{-\infty}^{s_2} f_n(s_1, s_2 \cdots, s_n) dB_{s_1} \cdots dB_{s_n} \\ &:= \int_{\mathbb{R}^n} f(t) dB_t^{\otimes n} \end{aligned}$$

Moreover, there is a isometry

$$E[F^2] = \sum_{n=0}^{\infty} n! \|f_n\|_{L_{(\mathbb{R}^n)}}^2$$

Itô integral can be expended by the Wiener-Itô chaos expansion. If F(t) is adapted, then

$$f_n(s_1, s_2, s_3 \cdots s_n, t) = 0 \;\; for \;\; s_i > t \tag{2.10}$$

Introduce $\hat{f}_n$ which is a symmetrization of f_n.

$$\hat{f}_n(s_1, s_2, s_3 \cdots s_n, t) = \frac{1}{n+1}(\overbrace{f(t, s_1, s_2, \cdots s_n) + f(s_1, t, s_2, \cdots s_n) \cdots}^{=0} + f(s_1, s_2, s_3 \cdots t))$$

Follow the property (2.10), a equation is given:

$$\hat{f}_n(s_1, s_2, s_3 \cdots s_n, t) = \frac{1}{n+1} f(s_1, s_2, s_3 \cdots t) \tag{2.11}$$

Then Itô-integral of $F(t)$ is now

$$\begin{aligned} &\int_{\mathbb{R}} F(\omega, t) dB(t) \\ &= \int_{\mathbb{R}} \sum_{n=0} I_n(f_n(s)) dB(t) \\ &\overset{2.11}{=} \int_{\mathbb{R}} \sum_{n=0} (n+1) I_n(\hat{f}_n(s)) dB(t) \\ &= \sum_{n=0} I_{n+1} \hat{f}_n \end{aligned} \tag{2.12}$$

Itô-integral to a adapted process $F(\omega, t)$ can be written as:

$$F(\omega, t) \overset{\text{Itô-integral}}{\rightarrow} \sum_{n=0} I_{n+1}(\hat{f}_n) \tag{2.13}$$

The Wiener-Itô chaos expansion can be expended to not adapted process like fBm. But before that, a rewritten Wiener-Itô chaos expansion is convenient.

Definition 2.6 (The Wiener-Itô chaos expansion theorem II) *Let $F \in L^2(\mu)$. Then there exists a unique family $\{c_\alpha\}_{\alpha \in \mathcal{J}}$ of constants $c_\alpha \in \mathbb{R}$ such that*

$$F(\omega) = \sum_{\alpha \in \mathcal{J}} c_\alpha \mathcal{H}_\alpha(\omega) convergence \;\; in \; L^2(\mu)$$

where

$$\mathcal{H}_\alpha(\omega) = h_{a_1}(< \omega, \xi_1 >) \cdots h_{a_n}(< \omega, \xi_n >)$$

h_n are Hermite polynomials and ξ_n are Hermite functions. $\mathcal{J}$ denote the set of all multi-indices $\alpha = (\alpha_1, \alpha_2, ...)$ of finite length, with $\alpha_i \in \mathbb{N} \cup 0 = 0, 1, 2, ...$ for all i. Moreover, the isometry is given as:

$$E[F^2] = \sum_{\alpha \in \mathcal{J}} c_\alpha^2 \alpha!$$

The chaos expansions for B_t and B_t^H are

$$\begin{aligned} B_t &= \langle \omega, I_{[0,t]}(\cdot) \rangle = \left\langle \omega, \sum_{k=1}^{\infty} (I_{[0,t]}, \xi_k)_{L^2(\mathbb{R})} \xi_k \right\rangle \\ &= \sum_{k=1}^{\infty} (I_{[0,t]}, \xi_k)_{L^2(\mathbb{R})} \langle \omega, \xi_k \rangle = \sum_{k=1}^{\infty} \int_0^t \xi_k(s) ds \mathcal{H}_{\epsilon^{(k)}}(\omega) \end{aligned} \tag{2.14}$$

To proof (2.14), the following results are needed:

$$\epsilon^{(k)} = (0, 0, \cdots, 1) \in \mathbb{R}^k$$

and

$$\mathcal{H}_{\epsilon^{(k)}}(\omega) = h_1(\langle \omega, \xi_k \rangle) = \langle \omega, \xi_k \rangle = \int_{\mathbb{R}} \xi_k(t) dB_t$$

B_t^H is calculated in the same way

$$B_t^H = \sum_{k=1}^{\infty} \int_0^t M\xi_k(s) ds \mathcal{H}_{\epsilon^{(k)}}(\omega) \tag{2.15}$$

Step 7

Definition 2.7 (White noise) *Definition of white noise W_t is*

$$W_t = \sum_{k=1}^{\infty} \xi_k(t) \mathcal{H}_{\epsilon^{(k)}}(\omega)$$

and fractional white noise W_t^H is

$$W_t^H = \sum_{k=1}^{\infty} M\xi_k(t) \mathcal{H}_{\epsilon^{(k)}}(\omega)$$

And the relation between W_t and W_t^H is obviously

$$W_t^H = MW_t$$

Step 8

Definition 2.8 (Wick product) *Define the Wick product for $F(\omega) = \sum_\alpha c_\alpha \mathcal{H}_\alpha(\omega)$ and $G(\omega) = \sum_\beta d_\beta \mathcal{H}_\beta(\omega)$, their Wick product $(F \diamond G)(\omega)$ is:*

$$(F \diamond G)(\omega) = \sum_{\alpha,\beta} c_\alpha d_\beta \mathcal{H}_{\alpha+\beta}(\omega)$$

Finally, the definition of WIS integral.

Definition 2.9 *Suppose $f(t, \omega) : \mathbb{R} \times \Omega \to \mathbb{R}$ is Skorohod integrable. Then*

$$\begin{aligned} \int_{\mathbb{R}} f(t, \omega) \delta B_t &= \int_{\mathbb{R}} f(t, \omega) \diamond W_t dt \\ \int_{\mathbb{R}} f(t, \omega) dB_t^H &= \int_{\mathbb{R}} f(t, \omega) \diamond W_t^H dt \end{aligned}$$

2.2.2 Fractional Itô formula

The important results for fractional Itô calculus are presented in this section. The first and most wildly used is Itô formula in fractional version.

Theorem 2.1 (A fractional Itô formula) *Let $H \in (0,1)$. Assume that f(s,x) : $\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ belongs to $C^{1,2}(\mathbb{R}\times\mathbb{R})$, and assume that the random variables*

$$f(t, B_t^H), \int_0^t \frac{\partial f}{\partial s}(s, B_s^H)ds \ \textit{and} \ \int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s^H)s^{2H-1}ds$$

all belong to $L^2(\mathbb{P})$. Then

$$\begin{aligned} f(t, B_t^H) = f(0,0) + \int_0^t \frac{\partial f}{\partial s}(s, B_s^H)ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s^H)dB_s^H \\ + H\int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s^H)s^{2H-1}ds \end{aligned}$$

A fractional Itô isometry is proofed, too.

Lemma 2.1 (Fractional Itô isometry) *If f, g belong to $L^2(\mathbb{R})$, then $\int_{\mathbb{R}} f_s dB_s^H$ and $\int_{\mathbb{R}} g_s dB_s^H$ are well defined zero mean, Gaussian random variables, then*

$$E\left[\int_{\mathbb{R}} f_s dB_s^H \int_{\mathbb{R}} g_t dB_t^H\right] = \int_{\mathbb{R}}\int_{\mathbb{R}} f_s g_t \phi(s,t) ds dt$$

where

$$\phi(s,t) = H(2H-1)|s-t|^{2H-2}, \textit{where } \ H \neq \frac{1}{2} \tag{2.16}$$

Refer to [4] for details and proofs of Theorem 2.1 and Lemma 2.1.

2.2.3 Ornstein-Uhlenbeck process

In this subsection, the results from previous subsections are used to solve a stochastic Partial differential equation(SPE). Solution of this type equation is called fractional mean-reversion Ornstein-Uhlenbeck process. This subsection is a example for using of fractional Itô formula. Another reason why the Ornstein-Uhlenbeak equation has a special place in this thesis is that temperature dynamic is described by the Ornstein-Uhlenbeak process in *chapter 4*, The solution of the Ornstein-Uhlenbeak equation is needed for pricing of temperature-based weather derivatives.

Definition 2.10 (Fractional Ornstein - Uhlenbeck processes) *The fractional mean reverting Ornstein - Uhlenbeck process is the solution X_t of the stochastic differential equation*

$$dX_t = \kappa_t(\theta_t - X_t)dt + \sigma_t dB_t^H, \ \ X_0 = x \tag{2.17}$$

where κ_t θ_t and σ_t are bounded deterministic functions. B_t^H is a fBm.

The equation (2.17) can be solve followed:

Proposition 2.1 *Let*

$$K_t = exp(-\int_0^t \kappa_s ds)$$

then the solution to the equation (2.17) is

$$X_t = xK_t + K_t \int_0^t \kappa_s \theta_s K_s^{-1} ds + K_t \int_0^t \sigma_s K_s^{-1} dB_s^H \tag{2.18}$$

and the distribution of X_t is given by:

$$X_t \sim N(xK_t + K_t \int_0^t \kappa_s \theta_s K_s^{-1} ds,\ \ K_t^2 \int_0^t \int_0^t \sigma_u \sigma_s K_u^{-1} K_s^{-1} \phi(u,s) duds)$$

where ϕ is given by the equation (2.16)

Proof.

$$\begin{aligned} d(exp(\int_0^t \kappa_s ds) X_t) &= \kappa_t K_t^{-1} X_t dt + K_t^{-1} dX_t && \text{(Fractional Itô formula)} \\ &= K_t^{-1}(\kappa_t X_t dt + \kappa_t(\theta_t - X_t)dt + \sigma_t dB_t^H) && \text{(by equation 2.17)} \\ &= K_t^{-1}(\kappa_t \theta_t dt + \sigma_t dB_t^H) \end{aligned}$$

Integrate both sides,

$$\begin{aligned} K_t^{-1} X_t &= x + \int_0^t K_s^{-1} \kappa_s \theta_s ds + \int_0^t K_s^{-1} \sigma_s dB_s^H \\ X_t &= xK_t + K_t \int_0^t \kappa_s \theta_s K_s^{-1} ds + K_t \int_0^t \sigma_s K_s^{-1} dB_s^H \end{aligned}$$

To find the distribution of X_t, the key is from *Definition 2.1.* The fBm is a Gaussian process as well as classical Bm and the only stochastic term in X_t is

$$\int_0^t \sigma_s K_s^{-1} dB_s^H \tag{2.19}$$

This term (2.19) is normal distributed for the same reason as in the case of classical Bm. Expectation of integral of fBm is zero as well as integral of Bm, because the increments of fBm are normal distributed with zero mean, too. To calculate variance of the term (2.19), fractional Itô isometry is useful.

$$\begin{aligned} Var(\int_0^t \sigma_s K_s^{-1} dB_s^H) &= E((\int_0^t \sigma_s K_s^{-1} dB_s^H)^2) && \text{(zero mean)} \\ &= \int_0^t \sigma_u \sigma_s K_u^{-1} K_s^{-1} \phi(u,s) duds && \text{(Fractional Itô isometry)} \end{aligned}$$

In order to proof *Proposition 2.1*, the only needs now are the basic properties of expectation and variance,

$$E[X_t] = xK_t + K_t \int_0^t \kappa_s \theta_s K_s^{-1} ds + 0$$

$$Var\left[X_t\right] = Var\left[K_t \int_0^t \sigma_s K_s^{-1} dB_s^H\right] = K_t^2 \int_0^t \sigma_u \sigma_s K_u^{-1} K_s^{-1} \phi(u,s) duds$$

Since the $(B_t^H)^2$ term does not involved in the fractional Ornstein-Uhlenbeck equation, the variance is the only different between the fractional and the classical Ornstein-Uhlenbeck process.

2.2.4 Girsanov theorem

In the end of this section, a fractional version of Girsanov theorem is given. Because of the relation

$$B^H(t) = \int_{\mathbb{R}} M_s dB_s$$

The classical Girsanov theorem can be applyed to fBm.

Theorem 2.2 (Fractional Girsanov theorem) *Let θ, Θ be measurable functions with support on [0,T], where θ is continuous and Θ is the solution of the integral equation $\int_{\mathbb{R}} \Theta_s \phi(s,t) ds = \theta(t)$, then*

$$\tilde{B}_t^H = B_t^H + \int_0^t \theta_s ds$$

is a fractional Brownian motion under the probability measure Q on $(\Omega, \mathcal{F}_T^H)$, which is equivalent with P and

$$\frac{dQ}{dP} = exp(-\int_{\mathbb{R}} \Theta_s dB_t^H - \frac{1}{2}\|\Theta\|_{L^2}^2)$$

being the Radon-Nikodym derivative.

The proof of fractional Girsanov theorem is vital for construction of a fractional finance market. The possibility of existents for a absolutely continuous risk-neutral probability measure $\mathcal{Q}$ to $\mathcal{P}$ is encouraging in hope of establishingS a arbitrage-free fractional market model.

Chapter 3

Weather Derivatives

3.1 The weather Derivatives market

Unpredictable elements in the finance world such as price, foreign exchange and the interest rate, have a common name: risk. The development of financial derivatives gives the tools to manage this category of unwanted risk. But what about another category of risk, such as weather? Unpredictable weather costs money. Many industries are affected by weather risk. Weather conditions like temperature, snow and rain fall, have significant influence on businesses and organizations. Traditionally, the weather risk is accepted as a fact, a risk the industries must take. But the idea of pricing mother nature and developing a instrument to manage weather risk is always discussed. With the participation from energy and insurance industries, a new asset class, so-called weather derivatives were born.

The weather derivatives market is a relatively new member of the finance market. In 1997, the first weather derivative was conceived and executed between three early pioneers in the market- energy traders Aquila, Enron,and Koch Industries. The first contracts were traded as Over-the-counter(OTC) derivatives. The market has grown rapidly and the Chicago Mercantile Exchange(CME) launches first standardized exchange weather derivatives in September 1999. In the year 2003-2004, the total limit of weather transactions executed amounted to $4.7 billion. In the period 2005-2006 this number jumped nearly tenfold to $45.2 billion.[1]. Today, CME offers weather products based on temperature index for 18 cities in U.S., and nine European and two Asia-Pacific cities. In the present market, most trading is still over-the-counter, standardized weather derivative contracts are now listed on the Chicago Mercantile Exchange (CME), the Intercontinental Exchange (ICE), and the London International Financial Futures and Options Exchange (LIFFE).

The weather derivatives market is such a complex, many fields are involved. Knowledge in meteorology, statistic, mathematic and finance are central. And many issues about the weather derivatives market have received a lot attention from academia. Completeness of the market, risk management and pricing approach are some of the most discussed themas. In this chapter, a brief intro-

[1]Number is according to homepage of WRMA

duction is given for the weather derivatives market. Structure of the market is peculiar, and include many elements. A simply example and a interesting real-life example are given, in order to illustrate the practical side of the weather market. In the end of the chapter, synopsis of four major approaches for pricing are given. There are two books attempt to describe the weather derivatives from top to bottom: Insurance and weather derivatives - From exotic options to exotic underlings, edited by Hèlyette Geman [11] and Weather derivative valuation by Stephen Jewson and Anders Brix[12].

3.1.1 Weather and weather exposure on business

That is a fact weather condition have influence on business. The influence can mean profits or losses, and it appears on every chain of the business, from production, transport to sales. In extreme cases, the results is catastrophic and in the other cases just small reductions in revenues. Catastrophe insurance futures contracts(CAT) are designed specially for losses caused by earthquake, extra-tropical storms and other nature catastrophes. The weather derivatives, however are not designed for catastrophic events. Non-catastrophic influence of weather can be warmer summer and colder winter than average, rainy and dry periods, long snow period and so on. It seems that the whole economy is potential participator in the weather derivatives market. There is however en missing link. The correlation between weather and losses must be significant and easily to be quantized. They electricity consume is closed related to temperature, and can be quantifies in degree days, therefor the electricity industries is the major participators from the early stage of the market.

The pay-off of the weather derivatives are often less correlated with any other financial instruments, which makes the weather derivatives an outstanding alternative financial strategic. The weather derivatives helps companies to lower volatility in profit. A low volatility is beneficial for a company in several ways: for borrow money from bank, for higher share price or for a more liquidity for cash flow.

The pay-off of weather derivatives are depends on a weather index, and it is unlikely that the pay-off will be the exactly amount of the losses. This fact is so-called basic risk and need to be studied closely for individuals interest. And this basic risk, is the fundamental difference for weather derivatives and weather insurance. To receive the pay-off, the holder of a weather derivative contract do not need to surfer losses caused by weather conditions. And therefore speculations on weather derivatives is allowed just as any other finance marked. For the primary participators of the market, they can hedge their weather risk. For banks, hedge funds and in some point of view, reinsurance companies, as speculators, they make extra money by their understanding of the market. To understand the market, the market structure is a place to begin.

3.1.2 Market structure

As the director of the weather derivatives group at Koch Industries claimed in a article [10], the motivation behind the weather contract is: Though one cannot physically change the weather, one can change one's weather exposure financially using the appropriate derivative instrument. Now the questions are: what is something special of the weather derivatives, how is the structure of weather derivatives?

Elements of weather derivatives

Weather Station

Weather contracts are linked to one or several specific weather stations. Most of the contracts are based on the observations from a single station, and there exist contracts that take a weighted sum over multiple stations. Stations that located at airports and large metropolitan areas are more popular. The most used measurement is daily temperature.

Contracts types

Weather derivatives use the common finance contracts. The most traded instruments are: put option, call option and swap. They are defined as follow

Definition 3.1 (Call and Put option) *Let S_t denote the underlying index and K is the strike. A call option is given by*

$$max(0, S_t - K) \text{ or } (0, S_t - K)^+$$

and a put option is given by

$$max(0, K - S_t) \text{ or } (0, K - S_t)^+$$

Call and put with the same strike are opposite to each other. The *figure A.1* is the pay-off of options.

Swaps are cash flows or other underlying index which two parties find mutually beneficial to switch between them.

All the three instruments can be arranged as future or forward contracts. In the OTC market, more exotic contracts are addressed, such as digitals, collars and barrier option. Readers may be confused of above mentioned types of contracts, although they are the same types in the traditional finance market. The following points describe the differences:

- An option gives the buyer right but not the obligation to buy or to sell the underlying index at a later time with an agreed price.

- A swap is an agreement between two parties to exchange future cash flows according to a prearranged formula. A swap involves always two pasties. A swap is often arranged as portfolios of forward or future contract.

- No price is paid at the beginning of a swap, therefore an exercise index must be chosen so that the expected payout will be zero.

- A future is exchange-traded, while a forward is traded over-the-counter. Thus a futures is standardized and faces an exchange, while forward are customized and faces a non-exchange counter party.

- A future is margined, while a forward is not. Thus a future has significantly less credit risk, and have different funding.

The above mentioned properties above are common for traditional finance derivatives and weather derivatives, but there are two major differences between them. The first is influence from the insurance industry. The payout has typically an upper limit. The second is underlying index.

Index

Alternatives of the underlying index depends on what sorts of weather risk one wishes to hedge. Temperature is a significant risk for mange industries, for example, the energy industry. In the meantime, road salt companies may prefer snow fall as the weather measurement. The next element is term. The choices of term depend on the nature of weather exposure but also are affected by the markets participants. The most common terms in the market are November 1 through March 31 for winter season contracts and May 1 through September 30 for summer contracts. As the market growing, one-month ever one-week contracts have received more and more attentions. More customized choices of terms are arranged in the OTC market, such as weight up weekends or some special days. Term operations are the third element of the underlying index. The most common operations are: sum, average, minimum and maximum. The operations are taking over the daily index over a period.

Three types of degree-days indexes are present here. There are numerous underlying indexes in the weather market, however these three most common are Heading degree days (HDD), Cooling degree days (CDD) and Cumulative average Temperature (CAT). HDD and CDD are primarily used by electricity companies. HDD is defined as follow:

Definition 3.2 (Heating degree days) *Let X_t be daily mean temperature. The heating degree days H_t is defined as follow*

$$H_t = (18^\circ - X_t)^+$$

HDD can trades over a period from T_1 to T_2, typically in winter months. The HDD over a period is defined as followed

$$H_t(T_1, T_2) = \int_{T_1}^{T_2} H_t dt \tag{3.1}$$

CDD is defined in simile way:

Definition 3.3 (Cooling degree days) *Let X_t be daily mean temperature. The cooling degree days C_t is defined as follow*

$$C_t = (X_t - 18^\circ)^+$$

and the CDD over a period is defined as

$$C_t(T_1, T_2) = \int_{T_1}^{T_2} C_t dt \tag{3.2}$$

The last underling index to be mentioned is

Definition 3.4 (Cumulative average Temperature) *The accumulated average temperatures over a period* $[T_1, T_2]$

$$\int_{T_1}^{T_2} X_t dt$$

These three indexes are available at CME. CME Weather contracts for the winter months in U.S. and the European cities are classified according to HDD values. The contracts for cities in U.S. in the summer months are geared to an index of CDD values. In Europe, CME Weather contracts for the summer months are based on CAT. Most of attention will be paid to this three index in rest of the thesis.

<u>Contract sizes</u>
The contract sizes are decided by the tick size and the limit. The tick of a weather derivative contract is that how much the pay-off changes per unit of the index. The tick is used to link the financial value of the contracts to the value of the underlying index, since the underlying index is a physical measurement and unexchangable. Most of the weather derivative contracts have the payout limited to a maximum (or minimum) possible value, known as the limit. This value is most commonly specified as a financial amount, but can also be specified in terms of the index value at which the limit is reached. In the OTC market, the tick and the limit are highly customized. But typically $5000 per degree day with a payment limit of $ 2 million. The standardized contract in CME has a tick of $ 20, and no payout limit. How ever a maximum order size is limited at 10 000 contracts, net long or net short in all contract monthly combined.

3.2 Example

The weather derivatives market is constructed by many usually uncorrelated factors, that makes the market difficult to understand. Two examples are given to illustrate the pay-off and application of the market.

3.2.1 An simply example

This example gives a pay-off function much like the call option on index HDD. The contract is signed as follow:

Table 3.1: *A Example for weather contract in OTC*

Contract Elements	Value
Weather station	Oslo, Blindern
Type	Call option
Strike	4500
Term	November 1,2009 to March 31, 2010
The underlying index	HDD
Contract size	Tick size = \$ 5000 per degree day
	Limit = \$ 2 million

The pay-off of the example is showed in figure A.2.

3.2.2 A real life example- Car insurance

In the early stage, participations of the weather derivative market are mostly from the energy industries. But recently the insurance industries, mostly non-life insurance, are paying more and more attention to the weather derivative market. The weather derivatives are considered as an alternative risk manage instruments, in addition to reinsurance. Since this is a cutting edge of risk management, many reaches are needed to establish the statistical relation between the weather condition and the potential losses. However, the statistical relation between weather and losses is often not straightforward and a conclusion is difficult to make.

A example is the car insurance. the empirical observations suggest a strong influence of weather to traffic accidents. The first snowfall in winter for example, causes always traffic chaos and many car clashes. A study by Tom Brijs et al.[7] on the effect of weather conditions on daily crash counts conform the empirical observations. A AR(1) is used to model the daily crash counts. Weather conditions, such as wind, temperature, sunshine, precipitation, air pressure and visibility are analyzed for a dataset of three big cities in the Netherlands(Utrecht, Dordrecht and Haarlemmermmeer),in the year of 2001. Temperature has found to be significant alone, without a combination with snowfall and rain. However, the relation of temperature and crash are not straightforward, an analysis is needed.

In this thesis, the studying subject is a dataset of daily losses reported to Gjensidige, one of the leading insurance groups in Nordic general insurance market. The attempt is to find the relation of the car losses and the temperature, specially for the group where daily temperature is under zero. The dataset includes daily losses of three Norwegian districts(Oslo, Bergen and Tromsø) from 2001 to 2008. The daily losses are classify into three categories: total (T), glass (G) and breakdown and damage to hull (D). The last to categories G and D should not have direct influence of temperature. A dataset of numbers of policies is also provided of Gjensidige. The figure A.3 show the numbers of policies and the figure A.4 are number of losses for both T and T-G-D, for the three districts. From the figures A.3 and A.4, it seems the number losses are positive related to number of policy. From the view of local difference, Oslo has the highest number of policy and therefor the highest number of losses. Tromsø is the opposite case. However, for the same location, the growth of policy over time is much higher than the growth of losses. Take Oslo as a example, the slop parameter for linear trend of policy is 2.34, and only 0.0015 for losses. In order to investigate the relation between the temperature and the losses, the losses T and T-G-D are normalized into number of losses per 10000 policies in following way [2]:

$$T = \frac{T}{\text{number of policy}} \times 10000$$

$$T - G - D = \frac{T - G - D}{\text{number of policy}} \times 10000$$

And the temperature data are divide into four groups:

- Group 1 - temperature above 20
- Group 2 - temperature between 10 and 20
- Group 3 - temperature between 0 and 10
- Group 4 - temperature under 0

The average of the total losses are plotted in the figure A.5. The figure indicate that the average of the total losses are higher in the cold period in all three districts. The pattern is most significant for Oslo. The T and T-G-D for Oslo is plotted in figure A.6, together with the average temperature of Oslo. The total losses are hither for colder days, but the T-G-D are nearly constant. The same pattern is repeated for Bergen and Tromsø. The averages of the losses are calculated for the four temperature groups as well. The result is illustrated by a boxplot. The figure A.7 is the boxplot of the 4 groups of temperature against the total losses. The Group 4 has a higher average than the other three groups, which conform the observation from the figure A.5.

The analysis performed in this section is basic. To argue the temperature-based weather derivatives as alternative risk management for the car insurance, many other factors have to be considered. Such as traffic level, types of car, snowfall and so on. Since the temperature is not the major factor for the car losses, the effect can not be concluded before removing the major factors.

3.3 Pricing a weather derivative

After the structure of weather derivative market is introduced, one of the most essential issue is pricing of weather derivatives. The practice is the buyer of a weather contract pays a price which is typically between 10% and 20% of the notional amount of the contract. As the market grows, more sofiscated pricing methods are under developing. There are several approaches for pricing. From

[2]Negative numbers of losses are replaced by zero. Reason can be doublet reported for glass and damage to hull

macroeconomic point of view, price level depends on supply and demand factors. Meanwhile, the weather market is an extension of traditional finance market, a pricing theorem based on stochastic analyses is a naturally suggestion. Furthermore the weather market carrys a mark from insurance market, therefor the actuarial pricing method should be considered. And for more exotic structures of derivatives, Monte Carlo pricing has a big advantage. A general description is given on all of four above mentioned methods. More details for the stochastic method are given in *chapter 4* and the Monte Carlo method are discussed in *chapter 6*.

3.3.1 The supply and demand method

One of the fundamental ideas behind the macroeconomic theorems is: the equilibrium of price level is given by the intersection of the aggregate supply curve and the aggregate demand curve. The labor market, the goods market and the financial market are all in equilibrium. Theoretically, this idea should also applies to the weather derivative market. The biggest difficulties are: liquidity of the market, the economic exposure to weather risk varying among industries and the untradable underling index.

The existence of a 'market price' decided by supply and demand, is determined by markets liquidity. Participating of the weather market is increasing. However a large number of contracts are traded in OTC market. The OTC market is highly customized. The contracts have a great range of variation, and liquidity in such market is low. In contrast, CME Weather markets are standardized and provide exceptional market liquidity and a dynamic trading by asset managers, multinational corporations, speculators, day traders, retail investors and investment banks. 5-10 locations in U.S. and 5 locations in Europe have sufficient liquidity in trading of HDD and CDD.

The supply and demand method has other issues than liquidity. The most critical is that the underlying index is untradable. That means buying and selling cannot influence the index directly. This fact ruins the foundation of the supply and demand method. In this case, the use of the tick to give the index a financial value can be a solution. However this solution just moved question from pricing the contracts to finding a tick value. Since the economic exposure to weather risk varies among buyers, a 'fair' tick value is unconceivable to be found. The supply and demand method is discussed in more details at [17].

3.3.2 The stochastic method

For traditionall options in the finance market, the Black-Scholes pricing model built on the arbitrage-free arguments, is employed. The Itô integral is the mathematical foundation of this model. To adapt the Black-Scholes model to the weather market, there are two challenges. The first one is that the weather market is incomplete market, as the underling index is untradable. Behavior of arbitrage opportunity on untradable asset need to be discussed. The second one is that, in this thesis, Brownian motion is repleted by fractional Brownian motion. The fBm is not semimartingale as mentioned in *chapter 2*. Semimartingale is one of the key properties Black-Scholes model requires. Using the stochastic method is in generally more complex than a actuarial method, but has advantages in accuracy and consistence for all underlying indexes on one location. This method is also called a analytical method.

3.3.3 The actuarial pricing method

The essence of the actuarial pricing method is the analysis of historical data in order to make an estimate of probability of future outcomes. More theoretical speaking, the method tries to derive the cumulative probability distribution functions(CDF) of the contracts pay-off. Technically, the method uses the historical data as time series of simulations to estimate parameters for the CDF, where the expectation μ and the standard deviation σ are two of the most important parameters. The most critical issue of this method is the historical data. At first, in average, fifty years historical data are available for each location. However due to autocorrelation among the historical data, the number of independent observations are mighty reduced. Beside that, a decision must be made about how mange years data are relevant. Over the time, climate are not necessarily constant. Actually, the global warming caused more extreme trends of climate, such as El Niño phenomenon. The trends vary both geographically and seasonally. As a results of the trends, the last 10 years data have more impact on estimation of the future outcomes. Based on this reason, the historical data may not be sufficient to prove the CDF. The Monte Carlo method has a great advantage in this point.

3.3.4 The Monte Carlo method

The Monte Carlo method is widely used in both finance and insurance. The method is an extension of the actuarial method. The benefit of this method is that the numbers of simulations are not limited any more. In Monte Carlo method, simulations are created via computer, based on a model fitted for the historical data. Estimation error can be eliminated by increasing the number of simulation, but the attention must be paied for model error. Still, for exotic structures of derivatives with the absence of analytical solution, and when historical data is insufficient, the Monte Carlo method is the only option.

Chapter 4

Pricing of weather derivatives in a fractional market model

In this chapter, the weather market will be considered as an extension of the traditional finance market, and the stochastic pricing method is used. The main motivation of developing a stochastic pricing method is probably a urge for accuracy. A stochastic model often has a great potential to achieve the goal. However between accuracy and potential accuracy, the model choice is vital. A model error has mainly to sources, model it self and implementation errors. A model error can cost much more then estimation errors, because any god efforts are all in wrong directions. A stochastic model is complex in general, and a stochastic model for the weather derivatives market which is incomplete faces more challenges. In addition to the choice of fBm as driving process, the model error is possibly one of major critic for employment of the stochastic model. The mathematics involve in the fBm and in construction of a fractional market model are no longer in the level for ordinary participators of the weather market. This is the second barrier for the stochastic model. These facts may slow down the applications of the stochastic model. The last issue is pricing formulas given in this chapter are difficult to calculate. The prices calculated by these formulas, represent an other economic interpretation. However, developing of fractional model has a great potential in accuracy. It needs to be tried and failed, by fitting model to its own output. It is a step forward for understanding the mechanism of the market and the understanding means possibly great economic profit.

The first try on the stochastic model, is the well known Black-Scholes model, which is celebrated by finance mathematicians. However the weather market is a typical example for incomplete market, and the employment of fBm as driving process, have consequences that the classical Black-Scholes model is no longer actually. A fractional version of the Black-Scholes is suggested.

The chapter is organized as followed: after description of the classical Black-Scholes model, arguments are given for the fractional Ornstein-Uhlenbeck as process for temperature, then the fractional Black-Scholes model is introduced, price formulas for HDD and CAT are present in the last section. In this thesis, focus is on temperature-based underlying index, since temperature is the most popular daily measurement of the underlying index. The three most traded underling indexes, HDD, CDD and CAT, are all based on daily temperature. The main references of this chapter are Benth [1], Brody et al. [8] and Biagini, Hu and Øksendal et al. [4].

4.1 The Black-Scholes model

Presume a market with two assets. A risky security such as stock (S_t) and a risk free asset suck as a bank account(R_t). the value of the stock is following a geometrical Brownian motion and the bank account has a constant interest rate r.

$$dS_t = \alpha S_t dt + \sigma S_t dB_t$$

$$dR_t = rR_t dt$$

A self-financing portfolio in this market has a value dynamic:

$$dV_t = a_t dR_t + b_t dS_t$$

Where (a_t, b_t) is an investment strategy. Two essential assumptions are:

i). The market is complete and arbitrage-free. This guarantees a risk-neutral probability measure $\mathcal{Q}$.

ii). The driven process is a Brownian motion

Under a risk-neutral probability measure $\mathcal{Q}$, the current value of all financial securities is equal to the expected value of the future pay-off of the securities discounted by the risk-free interest rate. The price dynamic in this market is given as a solution to the Black - Scholes partial differential equation:

$$C(t,x) = e^{-r(T-t)}\mathbb{E}[f(X_T(t,x))] \tag{4.1}$$

Where $X_T(t,x)$ is a value process of S_t under $\mathcal{Q}$ and f is the pay-off function. Using the property that B_t is normal distributed, $C(t,x)$ can be calculate. The result is known as the Black-Scholes formula.

An alternative approach is using the definition of risk-neutral probability measure. Let $\mathcal{F}_t$ be the σ-algebra generated by B_t, under $\mathcal{Q}$

$$V(t,x) = e^{-r(T-t)}\mathbb{E}_{\mathcal{Q}}[f(S_T(t,x))|\mathcal{F}_t] \tag{4.2}$$

And since there are no arbitrage opportunity, $V(t,x) = C(t,x)$. This is a more general approach.

The Black-Scholes formula can be extended to other underlying processes than the original geometrical Brownian motion, such as the Ornstein-Uhlenbeck process. The generalized Black - Scholes formula can be found in [15].

4.2 Dynamic of temperature

To price weather derivatives by using the stochastic method, a dynamic of temperature must first in place. Both empirical experience and statical studies suggest a fractional Ornstein - Uhlenbeck process to be the dynamic. This suggestion is based on two observations. Firstly, temperature is clearly local depended. At many locations, the historical data exhibit that temperature has a significant memory effect. A traditional Brownian motion has independent increments, and can not be sufficient for all locations. A fBm is therefor a more flexible driving process. By adjusting the Hurst coefficient of fBm, the model can be applied for locations worldwide. The next obvious trend of temperature has seasonal variation. Based on this trend, the mean-reverting Ornstein-Uhlenbeck stochastic process is proposed by Dornier and Querel (2000). In addition to a memory effect, a fractional version of Ornstein-Uhlenbeck process is suggested by Brody et al.[8] based on a analysis of the daily central England temperature (CET) recorded 1772 - 1999. In this thesis, the fractional Ornstein-Uhlenbeck process is employed. The process is given by

$$dX_t = \kappa_t(\theta_t - X_t)dt + \sigma_t dB_t^H, \;\; X_0 = x \tag{4.3}$$

X_t is the average temperature of day t, B_t^H is an fBm, κ_t is the mean-reversion rate, σ_t is the standard error of temperature and θ_t is the trend. The solution of (4.3) was given in *Chapter 2*, page 17, together with probability distribution of X_t. In the following *chapter 5*, κ_t, θ_t and σ_t will be estimated for a Norwegian dataset, for five districts.

The fractional Girsanov theorem gives transformation of X_t under $\mathcal{Q}$. Consider a Girsanov transform

$$\tilde{B}_t^H = \int_0^t \frac{\kappa_s}{\sigma_s} r_s ds + B_t^H$$

where λ_t is market price of risk, the dynamic of X_t under $\mathcal{Q}$ is given by

$$dX_t = \kappa_t(\theta_t + \lambda_t - X_t)dt + \sigma_t d\tilde{B}_t^H \tag{4.4}$$

and the solution of X_t under $\mathcal{Q}$ is

$$X_t = K_t(x + \int_0^t \kappa_s(\theta_s + \lambda_s)K_s^{-1}ds + \int_0^t \sigma_s K_s^{-1} d\tilde{B}_s^H) \tag{4.5}$$

In order to calculate CAT price, the cumulative temperature dynamic is needed as well. By using Fubini's theorem to exchange double integral of the stochastic part, $\int_0^T X_t dt$ under $\mathcal{Q}$ is given by

$$\begin{aligned}\int_0^T X_t dt = &\int_0^T K_t(x + \int_0^t \kappa_s(\theta_s + \lambda_s)K_s^{-1}ds)dt \\ &+ \int_0^T K_t \int_s^T \sigma_s K_s^{-1} d\tilde{B}_t^H\end{aligned} \tag{4.6}$$

4.3 The fractional Black-Scholes model

The Black - Scholes model cannot be directly applied to pricing of weather derivatives based on two major issues. In the first instance, the weather market is not complete as the underlying index is untradable. In an incomplete market, there exists at least two risk-neutral probability measures. Each of them can give an unlike value to $\mathbb{E}_Q[f]$. A price interval is given instead of one explicit price. Moreover, the fBm is not semimartingale, and the whole stochastic calculus building on the martingale property fall apart. To take the above mentioned challenges, Fred Espen Benth proposed a new definition of expectation, so called quasi-conditional expectation. Arbitrage-free price dynamics can be calculated for derivatives on temperature using quasi-conditional expectation. In this section, the definition of quasi-conditional expectation is given, together with a introduction of the structure for fractional Black-Scholes model. For special interests, recommendation of many useful theorems are presented in [1].

Definition 4.1 (The quasi-conditional expectation) *Let $X \in L^2(P)$ have Wiener-Itô chaos expansion I, $X = \sum_{n=0}^{\infty} I_n(f_n)$. The quasi-conditional expectation of X with respect to $\mathcal{F}_t^H$ is defined as*

$$\hat{E}\left[X|\mathcal{F}_t^H\right] = \sum_{n=0}^{\infty} I_n(f_n 1_{[0,t]}^{\otimes n})$$

where $\mathcal{F}_t^H$ is the σ-algebra generated by B_s^H for $s < t$.

The quasi-conditional expectation is an operator on the Wiener-Itô chaos expansion. This definition is quite abstract.The quasi-conditional expectation transfers many properties of traditional conditional expectation into fractional version. The transformation shares the same theoretical foundation like WIS integrals. Both use fractional white noise analysis to separate the fractional part of fBm into a new definition, and the calculation involved fBm can nearly remain the 'old fashion' on a new platform given by the new definitions. The follwing properties are vital:

- $\hat{E}\left[X|\mathcal{F}_0^H\right] = f_0 = E\left[X\right]$
- $\hat{E}\left[X|\mathcal{F}_t^H\right] = X$, If X is $\mathcal{F}_t^H$-adapted
- $\hat{E}\left[\hat{E}\left[X|\mathcal{F}_t^H\right]\right] = E\left[X\right], thelowofdoubleexpectation$

A new martingale type is defined by the quasi-conditional expectation.

Definition 4.2 (Quasi martingale) *If $M_t \in L^2(P)$ is an $\mathcal{F}_t^H$- adapted stochastic process. M_t is called a quasi-martingale if*

$$\hat{E}\left[M_t|\mathcal{F}_s^H\right] = M_s$$

for every $0 \leq s \leq t \leq \infty$

By the low of double expectation, the process $M_t := \hat{E}\left[X|\mathcal{F}_t^H\right]$ is a quasi-martingale whenever $X \in L^2(P)$. And the quasi-conditional expectation itself defines a quasi-martingale.

Moreover, a quasi-martingale representation theorem is given as the analog of the original. As known, Brownian motion is a martingale w.r.t the σ-algebras $\mathcal{F}_t$ generated by its own path, and so are many processes driven by Brownian motion. Part of purpose to quasi-martingale is to make a fractional Brownian motion quasi-martingale and extend useful properties of martingale into the fractional world.

Theorem 4.1 (The quasi-martingale representation theorem) *If M_t is a quasi-martingale with chaos expansion $M_t = \sum_{n=0}^{\infty} I_n(f_n 1_{[0,t]}^{\otimes n})$, then*

$$M_t = M_0 + \int_0^t N_s dB_s^H \tag{4.7}$$

where

$$N_t = \sum_{n=0}^{\infty} I_n((n+1)f_{n+1}(t,\cdot)1_{[0,t]}^{\otimes n})$$

Proof simply follows by the definition of fractional Itô integration which is defined similar as Itô integration. The sketch is given in *chapter 2*, page 14. In addition to a fractional Brownian motion, a number of stochastic process driven by fBm are quasi-martingale too. For pricing of weather derivatives, process of form is frequently involved.

$$Y_t = exp(\int_0^t a_s dB_s^H - \frac{1}{2}|a1_{[0,t]}|_\phi^2) \tag{4.8}$$

where the norm is defined by

$$|f,g|_\phi^2 : \int_{\mathbb{R}^n \times \mathbb{R}^n} f(s)g(t)\phi(s)\phi(t)dsdt$$

Fractional Itô formula and *Theorem 4.1* guarantees that Y_t is a quasi-martingale. In general, a stochastic process is quasi-martingale if it exists $\mathcal{F}_t^H$-adapted stochastic processes Y_t, Z_t such that

$$X_t = X_0 + \int_0^t Y_s ds + \int Z_s dB_s^H$$

Now together with the fractional Itô formula, the fractional Girsanov theorem and quasi-martingale, the ingredients for the fractional Black - Scholes model are all in place. Follow the same approach as the traditional Brownian motion, let $\mathcal{Q}$ be a risk-neutral probability measure and S_t is a quasi-martingale price process. Use the same portfolio strategy (a_t, b_t), a fractional Black -Scholes market is defined by

Definition 4.3 (A fractional Black - Scholes market/A WIS market model) *A fractional market has following elements*

A. To assets, a risk free bank account or bond which follows the dynamic

$$dR_t = rR_t dt, \quad R_0 = 1$$

A risky security S_t is given by

$$dS_t = \mu S_t dt + \sigma S_t dB_t^H, \quad S_0 = x$$

Both dynamics are $\mathcal{F}_t^H$ adapted.

B. A portfolio in the market is a measurable and $\mathcal{F}_t^H$ adapted stochastic process

$$\theta_t(\omega) = (a_t(\omega), b_t(\omega)), \quad 0 \leq t \leq T$$

C. The value at time t of a portfolio θ_t is defined by

$$V_t(\omega) = V_t^\theta(\omega) = V_0 + \int_0^t a_s(\omega) dR_s + \int_0^t b_s(\omega) dS_s$$

D. The portfolio θ_t is called self-financing if

$$dV_t(\omega) = a_t(\omega) dR_t + b_t(\omega) dS_t \tag{4.9}$$

To simplify, ω will not be mentioned always. A self-financing portfolio in the fractional market is also called a WIS self-financing portfolio since the integral w.r.t fBm is defined by Wick product. A portfolio θ in a WIS market model has no direct economic meaning. The advantage of the WIS market model is existence of arbitrage- free price dynamic. Assuming that the risky security S_t is the value dynamic of a weather derivatives on temperature, arguments are now given for an arbitrage-free price dynamic i the weather market.

Theorem 4.2 *Let Y be a contingent claim on temperature. Assume $\mathcal{Q}$ is an equivalent probability measure of $\mathcal{P}$ given by a Girsanov transform and $Y \in L^2(\mathcal{Q})$. An arbitrage-free price C_t of Y at time $t \in [0, T]$ is given by*

$$C_t = e^{-r(T-t)} \hat{E}_{\mathcal{Q}}[Y|\mathcal{F}_t^H] \tag{4.10}$$

where $\hat{E}_{\mathcal{Q}}$ is the quasi-conditional expectation under the probability $\mathcal{Q}$

Proof. In the traditional Black-Scholes market, if the discounted value process at time t $e^{-rt}V_t$is a martingale under $\mathcal{Q}$, the price dynamic is arbitrage-free. And in the fractional market, if $e^{-rt}V_t$ is a quasi-martingale under $\mathcal{Q}$, there is no arbitrage opportunity. Let

$$M_t := e^{-rT} \hat{E} - \mathcal{Q}[Y|\mathcal{F}_t^H] = e^{-rt} C_t$$

M_t is quasi-martingale by the low of double expectation. The discounted value process under $\mathcal{Q}$ is

$$d(e^{-rt}V_t) \overset{self-financing}{=} b_t d(e^{-rt}C_t) = b_t dM_t$$

Hence, $e^{-rt}V_t$ is a fractional quasi-martingale under $\mathcal{Q}$ and C_t is an arbitrage-free price dynamics.

Assume that $\mathcal{Q}$ is given by a Girsanov transform of $\mathcal{P}$, and the temperature dynamic under $\mathcal{Q}$ is given by the Girsanov transform as well. In order to calculate analytical price dynamic, a calculation technique for quasi-conditional expectation is needed. Since the definition of quasi-conditional expectation is abstract and based on chaos expansion, the solution directly from definition have nearly no practical meaning. Remind that the M operator transform integral of fBm into Bm, and the M operator is defined by Fourier transform. Using a Fourier approach to transform the abstract definition of quasi-conditional expectation into a more classical stochastic calculus solution, is therefor a natural choice.

Fred Espen Benth has completed the calculation in [1] and given following Proposition:

Proposition 4.1 *Let $b \in L^2_\phi(\mathbb{R})$, and define the function*

$$p_{t,T}(x) = \frac{1}{\sqrt{2\pi\Phi(t,T)}} exp(\frac{-x^2}{2\Phi(t,T)})$$

where $\Phi(t,T) = |b1_{[0,T]}|^2_\phi - |b1_{[0,t]}|^2_\phi$. Assume $f \in L^2(\mathbb{R}, p_{0,T(x)}dx)$, then

$$\hat{E}\left[f \int_0^T b_s dB^H_s | \mathcal{F}^H_t\right] = \int_{\mathbb{R}} f(y) p_{t,T}(\int_0^t b_s dB^H_s - y) dy \tag{4.11}$$

4.4 Pricing HDD and CAT

Finally, the pricing formula for HDD and CAT can be derived. First some notations is given

- T is the expiry time
- The term of HDD is given from T_1 to T_2
- C^{HDD}_t is the HDD price at time t and C^{CAT}_t is the CAT price at time t
- f is the pay-off function, $f^{HDD} = (18 - X_T)^+$ for HDD and $f^{CAT} = (18 - \int_0^T X_t)^+$ for CAT

Let the deterministic part of X_t be

$$\mathcal{D}_T = K_T(x + \int_0^T \kappa_s(\theta_s + r_s)K_s^{-1} ds)$$

and integrand of the stochastic part of X_t be

$$\mathcal{S}_{s,T} = \sigma_s K_T K_s^{-1}$$

then X_t under $\mathcal{Q}$ can be rewritten to

$$X_t = \mathcal{D}_T + \int_0^T \mathcal{S}_{s,T} d\tilde{B}^H_S$$

$$
\begin{aligned}
C_t^{HDD} &= e^{-r(T-t)}\hat{E}_{\mathcal{Q}}[f(X_T)|F_t^H] \\
&= e^{-r(T-t)}\hat{E}_{\mathcal{Q}}[f(\mathcal{D}_T + \int_0^T \mathcal{S}_{s,T} d\tilde{B}_s^H)|F_t^H] \\
&\overset{Proposition 4.1}{=} e^{-r(T-t)} \int_{\mathbb{R}} f(y) p_{t,T}(\mathcal{D}_T + \int_0^t \mathcal{S}_{s,T} d\tilde{B}_s^H - y) dy
\end{aligned}
\tag{4.12}
$$

The cumulative HDD over period T_1 to T_2 has pay-off function g given by

$$
\int_{T_1}^{T_2} (18 - X_u)^+ du = \int_{T_1}^{T_2} f^{HDD}(X_u) du \tag{4.13}
$$

Then the price dynamic for cumulative HDD C_t^{CHDD} is given by

$$
C_t^{CHDD} = e^{-r(T-t)}\hat{E}_{\mathcal{Q}}[\int_{T_1}^{T_2} f(X_u) du|F_t^H] \tag{4.14}
$$

To exchange the integral and the quasi-conditional expectation, the Fourier transform of function $\int_{T_1}^{T_2} f(X_u)du$ is needed. The Fourier transform will not be prensted in this thesis.

And the last price formula is for CAT. Since the dynamic of $\int_0^T X_t dt$ under $\mathcal{Q}$ is given by equation 4.6, we need only to redefine $\mathcal{D}_T$ and $\mathcal{S}_{s,T}$.

$$
\mathcal{D}_T = \int_0^T K_s(x + \int_0^s \kappa_u(\theta_u + r_u) K_u^{-1} du) ds
$$

$$
\mathcal{S}_{s,T} = \sigma_s \int_s^T K_T K_u^{-1} du
$$

Then the price dynamic of C_t^{CAT} is given by

$$
C_t^{CAT} = e^{-r(T-t)} \int_{\mathbb{R}} f(z) p_{t,T}(\mathcal{D}_T + \int_0^t \mathcal{S}_{s,T} d\tilde{B}_s^H - z) dz
$$

4.5 Comments on the fractional Black-Scholes model

In this chapter, a brief is given for the fractional Black-Scholes model and the analytical price formulas for weather derivatives HDD and CAT. Even though the whole framework for the fractional Black-Scholes model is an analog of the traditional Black-Scholes model, the solutions have quite different economic impacts. Two of the most significant differences in the fractional Black-Scholes model are

- The portfolio is not in the buy and hold fashion, since WIS integrals are based on Wick product, which cannot be directly linked to economic phenomenons.
- The price dynamic for HDD and CAT is now functions of t and T separately. The consequence is the contracts cannot be terminated or exchanged during the contract periods. Otherwise there will be arbitrage opportunity in the market.

The following Table 4.1 is an overview of the traditional and the fractional Black-Scholes model

Table 4.1: *Overview of the Black-Scholes model*

	Notation	Black-Scholes model	The fractional Black-Scholes model
Integral	$\int$	$\int X_s dB_s$ Riemann integral	$\int X_s dB_s^H = \int X_t \diamond W_s^H ds$ WIS integral based on Wick product
Asset process under $\mathcal{Q}$	dX_t	$dX_t = \kappa_t(\theta_t + r_t - X_t)dt + \sigma_t d\tilde{B}_t$	$dX_t = \kappa_t(\theta_t + r_t - X_t)dt + \sigma_t d\tilde{B}_t^H$
Portfolio	θ_t	(a_t, b_t) Buy and hold	(a_t, b_t) Macrocosm meaning? Total society utility?
Value process under $\mathcal{Q}$	dV_t	$dV_t = b_t dS_t$	$dV_t = b_t \diamond dS_t$
Self-finacing		Ja	Ja
Arbitrage		Arbitrage free	No strong arbitrage in WIS market
Price process	C_t	$C_t = e^{-r(T-t)} E_{\mathcal{Q}}(Y\|\mathcal{F}_t)$ Conditional expectation Semi-martingale $C_t(T-t, X_t)$	$C_t = e^{-r(T-t)} \hat{E}_{\mathcal{Q}}(Y\|\mathcal{F}_t^H)$ Quasi-conditional expectation Quasi-martingale $C_t(t, T, X_t)$

Chapter 5

Norwegian temperature data

The theoretical temperature dynamic is suggest in *chapter 4*. A empirical study of data is needed to conform the suggestion. And the parameters of the temperature dynamic must be estimated. The estimates are needed for pricing of the temperature-based weather derivatives, both in the analytical formulas and for the Monte Carlo method. A data analysis is a vital part of statistical studies. It is the connection of the mathematical model and the real-world. The quality of data analysis is a key factor for the accuracy of the proposed model.

The purpose of this chapter is to study the basic structure of Norwegian temperature data. The temperature data have a peculiar structure, which makes the method in the data analysis more complex. The temperature are analyzed in four stages: original(OR), detrended and deseasonalized (DD), residual (RES) and residual after removed the seasonal dependent sigma function (RES/SIGMA). The statistical and the fractional property are studied in all four stages. Many techniques for data analysis are used. A generator for fBm, together with three estimators of H values are introduced. The main reference for this chapter is Benth [2] and Brody et al.[8].

5.1 From the fractional Ornstein-Uhlenbeck to AR(1)

The fractional Ornstein-Uhlenbeck process is suggested as the dynamic of the temperature. In order to perform a analysis on the daily average temperature, a discrete model s needed. There are two reasons for preference of a discrete model. The first is the limitation of historical data. The dairy registration of temperature is the most complete data set which is available. The second reason is the underlying index is all written on degree-days, a discrete model is already sufficient. To derive a discrete version of the continuous-time Ornstein-Uhlenbeck model, recall first the definition of the fractional Ornstein-Uhlenbeck process in *chapter 4*

$$dX_t = \kappa_t(\theta_t - X_t)dt + \sigma_t dB_t^H \tag{5.1}$$

Where on day t, dX_t is the temperature change from day $t-1$ to day t, κ_t is the mean-reversion rate, θ_t is the trend and dB_t^H is an increment of fBm. However, the equation 5.1 is not reverting to θ_t in the long run. The θ_t is not a constant is the reason why the equation 5.1 failed to revert toward θ_t. This obsevition are suggested first by Dornier & Querul and is suggested in Benth [2] as well. An additional term is needed, to make the process really reverts to the θ_t. The term is of the form

$$\frac{d\theta_t}{dt} \tag{5.2}$$

The term 5.2 adjust the process to revert to θ_t in the long run, even θ_t is not a constant. The modified model for temperature is

$$dX_t = (\frac{d\theta_t}{dt} + \kappa_t(\theta_t - X_t))dt + \sigma_t dB_t^H \tag{5.3}$$

And the modified model can be solved by *Proposition 2.1*, the solution is

$$X_t = \theta_t + xK_t + K_t \int_0^t \kappa_s \theta_s K_s^{-1} ds + K_t \int_0^t \sigma_s K_s^{-1} dB_s^H \tag{5.4}$$

and

$$K_t = exp(-\int_0^t \kappa_s ds)$$

A discrete version of 5.3 is derived by following way. Let unit measurement of time be one-day, and $\Delta X_t = X_{t+1} - X_t$ denotes the change of temperature in one day, θ_t and dB_t^H are both rewritten in the discrete way as ΔX_t ,$\Delta t = 1$, the equation 5.3 can be written as:

$$\begin{aligned}
\Delta X_t &= \Delta\theta_t + \kappa_t(\theta_t - X_t)\Delta t + \sigma_t \Delta B_t^H \\
X_{t+1} - \theta_{t+1} &= (1+\kappa_t)(X_t - \theta_t) + \sigma_t \Delta B_t^H \\
X_{t+1} - \theta_{t+1} &\overset{properties\ iii) of fBm}{=} (1+\kappa_t)(X_t - \theta_t) + \sigma_t \epsilon_t
\end{aligned}$$

which is a AR(1) model with $\alpha = 1 + \kappa_t$ and $\epsilon_t \sim N(0,1)$. The analysis of Norwegian temperature data is based on the discrete time AR(1) model which is given as follow:

$$X_t - \theta_t = \alpha(X_{t-1} - \theta_{t-1}) + \sigma_t \epsilon_t, \ t = 0, 1, 2, \dots \tag{5.5}$$

Where on day t, X_t is the daily mean temperature,α is the autoregressive parameter, θ_t is the trend variable and ϵ_t is the noise. The model follow an autoregressive model of order 1 AR(1). The model is one of most employed in the temperature modeling.

In the following sections, an analysis is perform on the linear trend, the seasonality and the autoregressive property of the temperature data. The fractional property of the temperature is studied on OR, DD, RES and RES/SIGMA.

5.2 Statistical properties of the Norwegian temperature data

The datasets from five Norwegian districts, namely Kristiansand, Oslo, Bergen, Røros and Tromsø are fitted to the suggested model. The districts are located from the south to the north of Norway. Table 5.1 is a statistical overview of datasets. Since the five districts are geographical wildly spreaded, the differences among the mean, max, min and standard deviation of temperatures are expected. Values of skewness and kurtosis suggest that the empirical distribution of daily temperature is unsymmetrical. The results from the Jarque-Bera normality test conformed the suggestion. All these five districts are significant at 1 level for normal hypothesis. Having reranged the districts after Jarque Beta statistic, the empirical distribution of Bergen is the closest to the normal distribution and Røros is the opposite. In the rest of this chapter, ergen and Røros are used as two extreme examples, Oslo and Tromsø as two typical examples. To illustrate the non-normality, the histogram A.8 shows a clear sign of deviation from normal distribution. The red curves on the histogram are the best fitted normal distributions.

Table 5.1: *Daily mean temperature of 5 Norwegian districts*

Data	Daily mean temperature				
Series	6935 $\times$ 5 observations				
Period	From January 1, 1990 to December 31, 2008. February 29 is removed.				
Missing data	Replace by data from nearest station				
Units	Celsius degrees				
Jarque Bera [1]	Jarque-Bera is a normality test.				
	Kristiansand	Oslo	Bergen	Røros	Tromsø
Mean	7.64	6.8	8.34	1.29	3.28
Max	23.20	24.9	25.6	22	22.3
Min	-16.20	-18.2	-10.4	-36.2	-15.4
Std	6.76	7.91	5.67	9.06	6.48
Skewness	-0.20	-0.09	0.06	-0.63	0.11
Kurtosis	-0.50	-0.73	-0.51	0.48	-0.6
Jarque Bera	118.583	162.070	79.949	529.452	115.8066
P-value	0.000	0.000	0.000	0.000	0.000

[1] Jarque-Bera test is a goodness-of-fit measure of departure from normality, based on the sample kurtosis and skewness. The test statistic has an asymptotic chi-square distribution with 2 degrees of freedom, under the normal distribution hypothesis.

5.3 The linear trend and seasonality of the temperature

In this section, the trend variable θ_t is divided into two part:

- t_t - The linear trend
- s_t - The seasonal trend

The fact of the global warming and urbanisering, gives reasons to expect a positive linear trend of the temperature. The least squares estimate of the linear trend are given by the table 5.2. The linear trend is weak but does exist for each districts. The linear trend is weak statistically, due to the mean and standard deviation of daily temperatures. However the trend is significant in the climate context. The least squares estimate gives a 0.055° increasing of average temperature each year. For the last 19 years, the average increasing for the five Norwegian districts is 1.045 °C. Global near-surface temperatures have increased by 0.75° relative to the period 1860-1900, according to the instrumental temperature record. Compare this to numbers, the trend of global warming is strongly suggested.

Table 5.2: *Estimates of the linear trend*

Trend function	$t_t = b_0 + b_1 \times t$				
b_0	Intercept				
b_1	The least squares estimate				
R function	lsfit				
	Kristiansand	Oslo	Bergen	Røros	Tromsø
b_0	7.29535	6.04385	7.78237	0.84825	2.78520
b_1	0.00010	0.00022	0.00016	0.00013	0.000147

A closed look to histogram A.8, indicates a strong seasonal variation of the daily average temperatures. The figure A.9 certifies the seasonality and suggests that a seasonal function estimated by sinus/cosinus function. In other words, a Fourier series decomposition. The suggested decomposition is a Fourier series decomposition of order 1:

$$s_t = a_0 + a_1 \cos(\frac{2\pi}{365}(t - t_0)) \tag{5.6}$$

In table 5.3, the estimates is given for the Norwegian data series. Histogram for the detrended and deseasonalized temperature is displayed in figure A.10, together with the best fitted normal distributions. It's a clear improvement of normality of data series. And figure A.11 is included the average temperature, together with the estimated seasonal function and the detrended and deseasonalized temperature for Oslo. The rest of analysis in this thesis is based on the detrended and deseasonalized temperatures. This choice is not only made to justify the model 5.5, but also it is important for fractional property analysis. Data series with trend and seasonal behavior, appear to be non stationary. Non stationary series behave often like long memory processes. Models in this thesis, AR(1) in discrete time and Ornstein-Uhlenbeck process in continuous time are both stationary.

Table 5.3: *Estimates of the seasonality*

Seasonal function	$s_t = a_0 + a_1 \cos(\frac{2\pi}{365}(t - t_0))$				
a_0, a_1	Average level of detrended temperature				
t_0	Phase, is determined by the initial displacement at time t = 0				
R function	nlminb				
	Kristiansand	Oslo	Bergen	Røros	Tromsø
a_0	0.000000	0.00000	0.000001	0.00000	-0.000001
a_1	-8.322432	-10.00101	-6.739292	-10.51858	-7.665326
t_0	21.657315	17.05440	22.868789	17.62309	23.602160

5.4 Autoregressive parameter α

The discrete time model 5.5 is an AR(1) process. The parameter α indicates the speed of mean reverting. For $\alpha \geq 1$, the process is not stationary. And if $\alpha = 1$, then the process is a random walk. Theoretically, for $\alpha < 1$, the process is stationary, the mean and std of the process 5.5 is:

$$\begin{aligned} E(X_t - \theta_t - t_t|X_0) &= \alpha^t X_0 \overset{t\to\infty}{\to} 0 \\ Std(X_t - \theta_t - t_t|X_0) &= \sqrt{\frac{1-\alpha^{2t}}{1-\alpha^2}}\sigma \overset{t\to\infty}{\to} \frac{\sigma}{\sqrt{1-\alpha^2}} \\ Cov(X_t, X_{t-l}) &= \alpha^l Var(X_{t-l}) \overset{t\to\infty}{\to} \alpha^l \frac{\sigma}{\sqrt{1-\alpha^2}} \end{aligned} \tag{5.7}$$

The estimated α for the daily temperature is presented in table 5.4. Recall from

Table 5.4: *Estimates of α*

α	The autoregressive parameter				
σ^2	The estimated variance of process				
R function	ar				
	Kristiansand	Oslo	Bergen	Røros	Tromsø
α	0.7597	0.7864	0.7962	0.7219	0.7907
σ^2	4.663	4.744	3.432	12.75	4.667

the continuous-time Ornstein-Uhlenbeck model, the mean reverting parameter κ_t is a function of time. Stability of α over time should be studied. Fred Espen Benth et al. [2] performed an analysis of α and concluded that α does not depend on time or season for most Norwegian districts. α and κ are chosen to be constant in rest of thesis.

5.5 Residual analysis

Having detrended, deseasonalized and regressed Norwegian daily temperature, all deterministic parts of the model 5.5 are subtracted from the original data series. The residual is the only part which involves stochastic elements. Therefor the residual contains crucial informations for the dynamical model of temperature. The figure A.12 gives a general image of residuals. In fact, the empirical distribution of the residuals centered at zero and has much reduced standard deviation, but in the meantime, the skewness and the kurtosis are not much improved. Especially the kurtosis which s turned into positive for all 5 districts. High kurtosis suggests heavy tail for the empirical distribution. Normal distribution hypothesis are rejected at 1% significant level. Bergen is the most normal distributed and Røros is at the other end of the line. To reflect this, a simply table 5.5 as the table 5.1 has been made. The nest naturely question is about the stability of the empirical distribution through time. The qqpolt A.13 of the annual residuals for Røros demonstrates a consistence of the heavy tail property. The qqplot of rest of the districts indicates the parallel results.

Based on the residual analysis, two statistical properties of the residuals are concluded. The first one is that the residuals are not normal distributed. The assumption of the AR(1) model is often using normal distributed noise. The normal distribution fails to capture the heavy tail of the empirical distribution and therefor underestimated extreme high/low temperature. The information of the extreme value is critical in several fields, such as catastrophe insurance and climate changes. A Part of reason to retain normal distribution is, that the weather derivatives considered in this thesis is HDD, CHH and CAT. They are less sensitive to extreme value due to period tendency. The generalized hyperbolic distribution has been suggested by Fred Espen Benth et al.[2]. The distribution has five parameter and is flexible to fit the empirical distribution. The distribution is not suitable, since the fBm follows a normal distribution. And the model choice not only depends on accuracy to the historical data, but also needs to be robust, fast and easy to calibrate. Between the estimation errors of five parameter by the generalized hyperbolic distribution and the model error by normal distribution, the normal distribution is preferred. The second property is the strong local dependence of the residuals. And further more, the empirical distributions follow a pattern depending on the mean average temperature. The warm districts are in general more stable with low standard deviation and the empirical distribution are closer to the normal. On the contrary, the cold districts are often heavy tails. The summer is logical warmer then the winter, together with the pattern just suggested, a seasonal dependent standard deviation for the residuals may give a better estimation. This is discussed in the nest section.

5.6 Estimation of the σ_t

The last piece in this AR(1) model puzzle is the σ_t. As mentioned, a seasonal dependent standard deviation is proposed. This is one of the main findings in

Table 5.5: *Residuals of 5 Norwegian districts*

Data	Residuals of daily mean temperature				
Series	6935 $\times$ 5 observations				
Period	From January 1, 1990 to December 31, 2008. February 29 is removed.				
Jarque Bera	Jarque-Bera is a normality test.				
	Kristiansand	Oslo	Bergen	Røros	Tromsø
Mean	0.00	0.00	0.00	0.00	0.00
Max	12.86	11.07	7.64	19.60	9.82
Min	-10.67	-10.54	-7.55	-20.41	-9.07
Std	2.16	2.18	1.85	3.57	2.16
Skewness	0.20	0.05	0.21	-0.30	0.17
Kurtosis	1.63	1.10	0.38	3.68	0.60
Jarque Bera	814.493	351.053	92.047	4024.451	123.588
P-value	0.000	0.000	0.000	0.000	0.000

[3]. The same Fourier series decomposition technique is used for the estimation of the σ_t function. The function has the following form:

$$\sigma_t = c + \sum_{i=1}^{3} c_i sin(\frac{2i\pi t}{365}) + \sum_{j=1}^{3} d_j cos(\frac{2i\pi t}{365}) \tag{5.8}$$

The empirical σ is observed by resorting the data series into a matrix with dimension 19×365. Having taken the standard deviation of each column, the empirical $sigma_t$ is in a vector of length 365. The figure A.14 shows a clear seasonal variation of $sigma_t$, in addition to the local variation. The parameter of Fourier series decomposition is given in the table 5.6. The parameter c indicates the average of σ_t, it is reflected to the standard deviation of the residuals. The rest of the parameter is useful for Wavelet analysis. More details are discussed by Achilleas Zapranis et al in [22]. The figure A.15 includes the empirical and the estimate $sigma_t$ for Bergen and Røros.

5.7 Fractional analysis

After calibration of the AR(1) model, A analysis is performed on the fractional property of the Norwegian daily temperature. Three different estimation methods are performed on the temperature series of the following stages: original, detrended and deseasonalized, residual and residual after removing the seasonal dependent sigma function. The three methods employed in this thesis are: ST method, RS method and DFA method.

The Hurt coefficient is widely used as a measurement for the fractional property of random time series. There are excessive of methods to identify

Table 5.6: *Estimated parameters of σ_t*

Function	5.8
Std	Standard deviation of the residuals
Parameter	$c, c_1, c_2, c_3, d_1, d_2, d_3$
R function	nlminb

	Kristiansand	Oslo	Bergen	Røros	Tromsø
Std	2.16	2.18	1.85	3.57	2.16
c	2.00	2.06	1.75	3.20	2.07
c_1	-0.02	0.11	-0.06	0.35	0.05
c_2	-0.17	-0.06	-0.16	0.30	0.08
c_3	-0.04	0.02	0.00	-0.07	0.01
d_1	0.76	0.53	0.30	1.77	0.35
d_2	0.23	0.23	0.20	0.83	0.29
d_3	-0.06	-0.03	-0.01	0.01	0.01

the H value for a time series. Just to mention some, like spectral methods, maximum likelihood, time-scale methods and temporal methods. Jean-François Coeurjolly [9] gives a bibliographical overview of the estimation of fBm and H value. The methods in this thesis are anyway not included in [9]. The methods are chosen as they are already suggested by the studies related to temperature. To quantify the efficiency of the method, a generator for fBm is needed. Davis and Harte [21] suggested the Wood-Chan's method.

5.7.1 The Wood-Chan's method

The Wood-Chan's method is available for the whole Gaussian family. The basic idea is using autocovariance matrix. To simulate a random variable with dependence, the simplest way is to compute the autocovariance matrix M_{cov}, then to take square root of M_{cov} using such as Choleski decomposition, the last step is to multiple $\sqrt{M_{cov}}$ with an independent random variable. The major barrier is the computation time which is in order $\mathcal{O}(N^3)$. N is the number of simulation. The Wood-Chan's method is much improved in this matter by importing a circulant matrix C. In a circulant matrix, each row vector is rotated one element to the right which is the preceding row vector. Let $cov()$ be the autocovariance function, C is defined by:

$$C \begin{pmatrix} c_0 & c_1 & \dots & c_{m-1} \\ c_{m-1} & c_0 & \dots & c_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \dots & c_0 \end{pmatrix} \text{where} \quad c_j = \begin{cases} cov(j) & \text{if} \quad 0 \leq j < \frac{m}{2} \\ cov(m-j) & \text{if} \quad \frac{m}{2} < j < m-1 \end{cases}$$

Since C is defined symmetric and circulant, the first row has all theinformation of the whole matrix. A useful linear algebra result gives an unitary decomposition $C = Q\Lambda Q$, Λ is the diagonal matrix of eigenvectors of C, and Q can be decomposed by Fourier transform. Now if $Y = Q\Lambda^{\frac{1}{2}}Q * Z$, where Z is a random variable which is standard normal distributed with length m, then $Y \sim N(0, C)$ and the cumulative sum of $Y = (Y_0, \ldots Y_{N-1})$ is a fBm. More details can be found in [9]. The Wood-Chan's method is fast and efficient for the purpose like Monte Carlo simulation. The figure A.16 gives a idea on how fBm varies with different H values. The H values which presented in the figure are:$H = 0.1$, $H = 0.5$ and $H = 0.9$.

5.7.2 The ST method

The ST method is first introduced by Syroka and Toumi (2002) and implemented by Brody et al. [8] and Fred Espen Benth et al. [2]. The method is simple and efficient. For a time series X_t of length T, divide the series into N non-overlapping bins of length L. And computing the test statistic f(L)

$$f(L) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{\bar{X}_L^2 - \bar{X}}{sd(X)}} \tag{5.9}$$

where $\bar{X}_L$ is the average of each bins, $\bar{X}$ is the total average and $sd(X)$ is the total variance. It is known for a complete random and uncorrelated series:

$$f(L) \sim L^{\frac{1}{2}} \tag{5.10}$$

and for a correlated series: $f(L) \sim L^{-(1-H)}$. Based on this observation, H value can be estimated by compute f(L) for different value of L and H value is the slope coefficient of the linear regression.

5.7.3 The RS method

The rescaled range (RS) method is the original method to identify H value, it is introduced by Hurst(1951). The method has many modified version, such as Lo's modified RS(1991). Dividing the series in the same way as the ST method, for each bins, the original statistic duo to Hurst is defined as

$$\begin{aligned} X(t, L) &= \sum_{i=1}^{t}(X_i - \bar{X}_L), \quad \bar{X}_L \text{ is the average of each bins} \\ R(L) &= maxX(t, L) - minX(t, L), \quad \text{for} 1 \le t \le L \\ S(L) &= \sqrt{\frac{1}{L}\sum_{i=1}^{L}(X_i - \bar{X}_L^2)} \end{aligned} \tag{5.11}$$

Hurst observed an empirical low

$$R/S \sim (\alpha L)^H \tag{5.12}$$

The computation is similar to the ST method. The original RS method is sensitive to L, but not to H value. A Monte Carlo calibration is needed to order to find the optimal min and max value of L. Fotini Pallikari and Emil Boller [20] carried out an analysis of efficiency of the RS method.

5.7.4 The DFA method

The detrended fluctuation analysis (DFA) has an increasing popularity recently. The major advance is the DFA method can be applied directly to non-stationary series. The DFA method is a modified version of the ST method and the RS method in some sense. Again, X_t is divided into N non-overlapping bins with length L. The DFA statistic is defined as

$$\begin{aligned} f_i(t) &= a_i + b_i t \quad \text{the trend function i order 1} \\ z_t &= \sum_{i=1}^{t} X_i - f_i(t) \\ F(t, L) &= \sqrt{\frac{1}{L}\sum_{t=1}^{L}(z_t)^2} \end{aligned} \tag{5.13}$$

The DFA statistic follows the same empirical low as the RS statistic

$$F(t, L) \sim L^H \tag{5.14}$$

The DFA method can be applied directly to non-stationary series, as the local trend is estimated and subtracted for each bins. This detrending method is more efficient than subtract the least square trend for the whole series. Difference kinds of trends can be considered by the justification of the tend function $f_i(t)$, such as higher order of polynomial, sinus and exponential functions. For fair comparison to the other to methods, we choose the $f_i(t)$ in order 1. As same as the ST and RS method, the DFA method can be influenced by choice of L. Sebastian Michalski [18] studied the optimal length of bins, using Monte Carlo simulation and different fBm generator. The DFA method has been performed on a temperature records from Australia [13]. The finding about H value and geographic dependence is inspirational.

5.7.5 Monte Carlo calibration for the ST, RS and DFA method

A comparison of bias and standard deviation for the three methods is beneficial for the further analysis in this thesis. The Norwegian daily average temperature contains 6935 days for each districts, and all three methods are sensitive to the length of the series and length of the bins. We prefer a Monte Carlo simulation of the same data length. The table 5.7 presents Monte Carlo results for 10000 simulation of fBm increments generated by Wood-Chan's methed. The length of each fBm is 7000. The comparison is performed on different H values. The figure A.17 plots bias of the H value estimated by the three methods and the standard errors. The figure shows clearly that the ST method has the smallest bias and the RS method has the smallest standard errors. When the DFA method has a constant bias for 0.03, the ST method has a trend for underestimating for H value high then 0.7. If subtracting the constant bias 0.03 from the DFA method, DFA will be an outstanding estimator for H value. The RS method, even though it has the smallest standard error, overestimates for H value lower then 0.7, and underestimates the rest. The last issue is computing times for the methods. The ST method takes 0.23 sec for a single estimation, the RS method take 0.28 sec and the DFA method tops the record with 2,60 sec. Using The DFA method for already detrended time series is kind of waste the time. Recall that the DFA method here removes only linear trend, which is small with respect to the seasonal. With 'The sliding window'-technique, the DFA method is able to remove the seasonal trend [13].

Table 5.7: *Comparison of bias and standard deviation for the ST, RS, DFA methods*

H	H value								
Bias	Estimated average minus H								
Std	Standard deviation								
R function	woodFBM, ST, RS, DFA								
	$H = 0.1$	$H = 0.2$	$H = 0.3$	$H = 0.4$	$H = 0.5$	$H = 0.6$	$H = 0.7$	$H = 0.8$	$H = 0.9$
	Bias								
ST	-0.0005	0.0009	0.0015	0.0002	-0.0014	-0.0044	-0.0111	-0.0251	-0.0502
RS	0.1382	0.1106	0.0832	0.0571	0.0323	0.0066	-0.0219	-0.0560	-0.0992
DFA	0.0318	0.0285	0.0272	0.0273	0.0279	0.0302	0.0309	0.0331	0.0347
	Std								
ST	0.0175	0.0200	0.0219	0.0230	0.0232	0.0240	0.0245	0.0249	0.0244
RS	0.0071	0.0092	0.0112	0.0131	0.0146	0.0161	0.0169	0.0170	0.0168
DFA	0.0089	0.0127	0.0158	0.0187	0.0208	0.0235	0.0241	0.0242	0.0246

5.7.6 Comparison on fractional property for temperature data

The empirical experience suggests a long memory behavior for temperature. This observation motivated many studies to capture this property of the temperature. The two major branches are regression model such as AR(1) and fractional model based on fBm. Brody et al. [8] supposed a fractional model and estimated the H value for daily temperature in central England from 1772 to 1999 to be 0.61. And different kinds of regression models are wildly used. These two branches can also be combined. A comparison for OR, DD, RES and RES/SIGMA, gives a better understanding on fractional property for temperature data. The tabel 5.8 presents the H value for the four types data series, estimated by the ST, RS and DFA method, for the five Norwegian districts. For the RS method and the DFA method, some of the H values are bigger than 1 for the original temperature. This causes by a weakness of the methods. H values are defined in interval (0,1). The figure A.18 shows the plot of three method for the detrended and deseasonalized temperature for Oslo.

Considering the table 5.8 together with the table 5.7, some patterns can be found. The OR have high H value because the data series are not stationary. H value for the DD are lower, but still significant. Brody et al. [8] estimated the H value to be 0.61 for deseasonalized temperature of central England. H values for Norwegian districts are higher then England. A looking at the residuals, conform

Table 5.8: *H value for the four types data series, estimated by the ST, RS and DFA method, for the five Norwegian districts*

H	H value				
OR	The original temperature				
DD	The detrended and deseasonalized temperature				
RES	The residual after regression				
RES/SIGMA	The residual after removed the seasonal dependent sigma function				
R function	woodFBM, ST, RS, DFA				
	Kristiansand	Oslo	Bergen	Røros	Tromsø
	H value, ST				
OR	0.957	0.961	0.947	0.943	0.945
DD	0.755	0.734	0.714	0.663	0.718
RES	0.632	0.592	0.553	0.546	0.569
RES/SIGMA	0.639	0.598	0.558	0.575	0.572
	H value, RS				
OR	1.007	1.008	0.991	0.982	0.988
DD	0.793	0.769	0.759	0.727	0.779
RES	0.616	0.579	0.554	0.567	0.574
RES/SIGMA	0.610	0.574	0.559	0.580	0.582
	H value, DFA				
OR	1.064	1.036	1.075	1.066	1.049
DD	0.898	0.822	0.872	0.830	0.877
RES	0.619	0.577	0.551	0.593	0.571
RES/SIGMA	0.624	0.580	0.555	0.609	0.575

that the AR(1) regression is succeeded with removing fractional property. The H values for the residuals is reduced to under 0.6 for four of the districts. Dividing the standard errors has no significant effect on H values. The figrue A.19 shows the autocorrelation for the four stages of temperature series. The district used in the figure is Oslo. The seasonality of autocorrelation in the original series is removed in the detrended and deseasonalized series, and the autocorrelation is reduced toward zero in residuals. But some seasonality is still remained, suggest by a close look at the square residuals. The estimated sigma is succeeded to rewove this effect, as the last plot show. Even though patterns are consistent, the estimates are still geographical dependent. The estimates are actually sensitive to the length of data series too. The sensitivity to data length partly inherited from H value estimator, and is partly a property of the data series self.

As the results show, the AR(1) model is efficient to capture the long-rang dependence, the natural thought is: Can AR(2) do a better job? Again, using Oslo as an example, H values are estimated by the ST method for AR model in higher order. The table 5.9 presents the results. It is unexpected that the AR(1) model is the most efferent to capture the long-rang dependence of the data series and the fractional property in the residuals is not able to remove by the higher order AR model.

Table 5.9: *H values of the residuals for AR model*

District	Oslo			
H	H value			
RES	The residual after regression			
R function	AR, ST			
	$AR(1)$	$AR(2)$	$AR(5)$	$AR(10)$
Res	0.592	0.605	0.603	0.605

In the next chapter, the estimates are used to price the indexes of HDD, CDD and CAT and the effect on price by varying H values will be studied.

Chapter 6

Monte Carlo simulation

The Monte Carlo simulation is a widely used as pricing method. This method is more flexible than the analytical method. Even though the Monte Carlo method does not need many of mathematical calculations, still the temperature dynamic must in place. The estimates for the temperature dynamic are vital for accuracy of the Monte Carlo method. The Monte Carlo method it self is very robust and the error caused by the simulation are small in the respect to the model error. The accuracy of the Monte Carlo method depend on the whole framework, temperature dynamic and data analysis.

In this chapter, the Monte Carlo simulation technique is used to calculate the price of degree-days options, such as HDD, CDD and CAT. The chosen weather station is Oslo and parameters for the weather dynamic are estimated in the *chapter 5*. At the end of the chapter, a comparison is given for the prices driven by fBm with different H values.

6.1 Assumption

The Monte Carlo simulation technique is a way to numerically calculate the expected value $E_{\mathcal{Q}}(f(X_t)|\mathcal{F}_t)$, where X_t is the dynamic of temperature under risk-neutral probability measure $\mathcal{Q}$.

Two simplification are made for applying the Monte Carlo method. The first is the definition of prices. In the Monte Carlo method, the prices are defined as the discounted expectation of the value processes. This is however not fully consistent for the fractional Black-sholes model. The prices in the fractional Black-sholes model are defined via quasi-conditional expectation. Since there are no obvious way to simulate the quai-conditional expectation, and even if the quasi-conditional expectation is used, the prices will have an other economic interpretation. Therefor, this difference is just ignored. The second simplification is about the market price of the risk. To be able to simulate temperature under the risk neutral measure $\mathcal{Q}$, the market price of risk, λ_t, has be estimated . Since the liquidity of the weather market strongly depends on location, and there are no operative weather market for Oslo, the market price of risk is chosen to be zero in this thesis. The prices given in this thesis are therefor predicted price rather then future price.

6.2 Contracts on HDD, CDD and CAT

Prices for 6 contracts are simulated. They are written on call and put options for each HDD, CDD and CAT. This 6 combinations of the contract types and the underling indexes, are most common traded. For simplify, the tick size is chosen to be 1 NOR and the annual interest rate is chosen to be 5%. The details of contracts are presented in table 6.1. Only monthly contract are considered.

Table 6.1: *6 contracts of degree-days options*

Period	HDD is most traded in winter, CDD is most traded in summer					
	CAT is most trade in summer, for European cities					
Strike	Calculated by temperature 2008					
	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
Weather station	Oslo Blindern	Oslo Blindern	Oslo Blindern	Oslo Blindern	Oslo Blindern	Oslo Blindern
Index	HDD	HDD	CDD	CDD	CAT	CAT
Type	Call	Put	Call	Put	Call	Put
Period	January 2009	February 2009	June 2009	July 2009	June 2009	July 2009
Strike	620	560	12	21	479	564
Tick	1 NOK/HDD	1 NOK/HDD	1 NOK/CDD	1 NOK/CDD	1 NOK/CAT	1 NOK/CAT

As mentioned in the *chapter 3*, the HDD season is from November 1 through March 31, and the CDD season is from May 1 through September 30. In Europe, the CAT is traded in the same season as the CDD. The terms of the 6 contracts, are chosen to be the two coldest months for the HDD and the warmest months for the CDD and CAT. The choice of the terms are made to capture the max difference between the indexes. The CDD and the CAT are written on the same month for the same type of option, this makes the two indexes comparable. The last but important element of the contacts are the strike. As the options written on traditional asset, the strike determines the price. However, a stock price may double it self in anytime, but temperature of Oslo will not raise to $20°$ or down to $-30°$ in January. Then the degree-days are limited in some certain intervals. A strike out of the intervals, is not rational. And since the strong local dependence and seasonal variation of the temperature, the strike must be chosen carefully. The strikes of the 6 contacts are calculated based on temperature of 2008. The values of strikes are simply the HDD, CDD and CAT indexes calculated for the respect months in 2008.

6.3 Results

The simulations give following results in table 6.2. The indexes of the HDD and the CAT are closer to each others, while the indexes of the CDD are much lower. In North Europe, specially Norway, there are not many summers days the average temperature is higher than $18°$. This fact makes the CDD index low. A low index makes extra challenge for pricing. The prices will be much more sensitive to the strike. This is a possible reason why the CDD is not often traded on European cities.

The prices are stable for repeated simulations. This indicates a good temperature models. Compare the simulated indexes to the indexes calculated for the historical data, the values of the indexes are rational. It is difficult to calibrate the model for the Monte Carlo method, since there are no weather derivatives market operative for Oslo. However, the justifications for the model are not complicated, and can be easily done if the market data of prices are available.

Table 6.2: *Results of the 6 contracts*

T	1 mnd					
t	Day before the trading period					
Sim	Number of simulation = 100000					
Strike	Calculated by temperature 2008					
	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
Index	670	563.83	6.97	9.61	459	495
Price	39.17	4.029	0.43	9.06	0.11	53.47

6.4 Effect by varying the H values

In order to evaluate the effect by variation of the H values, the Option 1 are computed for different H values from 0.5 to 0.9. Almost all studies of temperatures are agreed that the temperature dynamic is persistent, therefor the simulation is focused on H values bigger than 0.5. The simulation results are shown in table 6.3. There are no difference neither indexes or prices for Option 1. The explanation of this result is visibly in the figure A.20. The figure gives plots of Oslo temperature with three H values, the estimated H value by the ST method, $H = 0.598$, $H = 0.5$ and $H = 0.9$. In the first plot, the fat black line is the registered temperature for Oslo, 2008, from the weather station Oslo, Blindern. The simulations with $H = 0.5$ are more concentrated, and $H = 0.9$ gives more dispersed simulations, without changing the expectation. Recall that the HDD strike is 18°, and the temperature in January in Oslo is around -3 degree, therefor the HDD strike have no effect, together with the fact that H values change only corrvariance, but not expectation of fBm, prices of the Option 1 remain the same for all H values. The same conclusion can be made for CDD and CAT as well.

Table 6.3: *Prices for option 1 with different H values*

Period	January				
T	1 mnd				
t	Day before trading period				
Sim	Number of simulation = 100000				
Strike	Calculated by temperature 2008				
	$H = 0.5$	$H = 0.6$	$H = 0.7$	$H = 0.8$	$H = 0.9$
Index	670.048680	670.048748	670.048675	670.048652	670.048623
Price	45.118914	45.143796	45.107263	45.213002	45.142108

However the H values do have effect on option prices. If the number of days which the daily average temperature is under zero, is considered as a option named with Froze. The table 6.4 gives prices for different H values for en such Froze option. The effect of H values is significant. From the estimates value, $H = 0.598$ to the classical Brownian motion with $H = 0.5$, the index changes 0.9 degree day and price 27.74 % . And from $H = 0.5$ to $H = 0.9$, the index changes 3.3 degree days and price 74.37%. The percent is using value of $H = 0.5$ as reference. There are logical explanation for this dramatic behavior of the prices. The index of HDD is decided by the expectation of the temperature alone, since the HDD strike at 18° is unreachable for a January month and H values have no effect on the expectation. The Froze index in the other hand, has a strike in 0°. This value is around the expectation of the January temperature. The higher H values make pathes of the simulations more dispersed, therefor the probability of pathes cross the 0° will be higher. This increasing probability make the expectation of the Froze index lower. It is easer to understand if the dispersion of pathes is explained in the following way: It makes a warm day to be warmer and a cold day to be colder. The first case reduce the Froze index and the second case changes nothing for the Froze index. Put the two cases together, the Froze index reduces for High H values.

From the comparison of the HDD and the Froz indexes, the conclusion is that the effect of the H values on price depends on choice of the underlying indexes. For indexes which have strike around the average temperature of the chosen term, the H values are significant factors for prices. The influence of H values is expected to be positive related to the length of the term, too. The studies of H values are essential for pricing the temperature-based derivatives.

Table 6.4: *Prices for Froze option with different H values*

Period	January				
Option	Froze, CALL				
Index	Numbers of icing day, Oslo				
T	1 mnd				
t	Day before trading period				
Sim	Number of simulation = 100000				
Strike	24				
	$H = 0.5$	$H = 0.6$	$H = 0.7$	$H = 0.8$	$H = 0.9$
Index	26.0266	25.1025	24.2434	23.4562	22.7210
Price	1.7866	1.2910	0.9187	0.6495	0.4579

Chapter 7

Conclusion and future research

7.1 Conclusion

The two main subjects being studied in this thesis are the fractional Brownian motion and the weather derivatives market, and the connections between these two fields. The subjects are presented for both the theoretical interest and applications.

A theoretical framework of the fBm and the weather derivatives market adapted to a fractional Black-Scholes model is introduced in the first part of this thesis. This part covers the first three chapters. The choice of using the WIS integrals is for the purpose of establishing an arbitrage-free fractional market model. The fractional Black-Scholes model is introduced as an analog of the classical Black-Scholes market. From the process of the underlying asset to the idea of the risk-neutral probability measure, from the well-known theorems such as Girsanov theorem and martingale representation theorem to the concept of conditional expectation and semimartingale, these are all extended to a fractional version. The surface of the fractional Black-Scholes model is much similar to the classical ones. However the theoretical foundation for the model is fractional white noise analysis, which differs from the applied approach for the construction of the classical Black-Scholes. As a direct consequence, the economic interpretations of the model are changed, for both portfolios and prices. According to Biagini,Hu and Øksendal [4], the meaning of values in a fractional market may move from a microeconomic to a macroeconomic view. However, using the fractional Ornstein-Uhlenbeck process as temperature dynamic, price formulas are derived for the most traded weather index HDD, CDD and CAT.

The weather derivatives market is introduced in *chapter 3* together with a real-life example in the car insurance. The data conformed that in the colder periods, there are more losses reported, number of losses is especially high for the period which daily temperature under zero. However, the temperature is not a major factor for the car losses, and conclusion cannot be taken before removing effect of other factors.

In the theoretical part, there is an attempt for accuracy of mathematical model to real-life, and the main motivation is to capture the long-rang dependence. A data analysis is performed on daily average temperature from five Norwegian districts to conform the suggested fractional Ornstein-Uhlenbeck process. A discrete version of the fractional Ornstein-Uhlenbeck process is employed in the analysis. There are two reasons for the preference of a discrete model. The first one is the limitation of historical data. The daily registration of temperature is the most complete dataset which is available for all the five locations. The second reason is that the underlying index is all written on degree days, and a discrete model is already sufficient. The AR(1) model is introduced as the discrete version. The AR(1) is fitted for all five locations, and the results conformed the empirical suggestion. The following unified patterns are observed in temperature data for all the five districts.

- The linear trend is not strong i statistical context.
- The seasonal trend is obviously.
- The autoregressive parameter α is round 0.75 for all the five districts and is stable over time.
- The original temperatures and the residuals are not normal distributed. A symmetric heavy tail distribution fit the residuals better.
- The seasonality in the variance of the residual is conformed. The variation is generally higher i cold period than warm period.
- The fractional property is conformed, even though the local dependence exists for estimates.
- The fractional property is strongly reduced from the detrended and deseasonalized temperatures to residuals, but is not able to be completely removed.
- AR(1) model is the most efficient in order to reduce the long-rang dependence. The hither order of AR model has no obvious improving.

In addition, the analysis suggests a local dependence of the estimates.

In order to estimate H values, a generator for fBm and three different estimators, the ST, RS and DFA method are introduced. A Monte Carlo calibration is performed. The ST method is the best, if combined with efficiency and accuracy. The DFA method has the best potential.

The last *chapter 6*, the Monte Carlo simulation is used for pricing contracts written on HDD, CDD and CAT. The market price of risk is not considered since the weather marked is not operational in Norway. The price levels are different for HDD, CDD and CAT, where the levels of HDD and CDD are closer. The possible reason is that temperatures in the North Europe are often lower than 18° even in summer, CDD has therefore a low index. The 18° may not be suitable for the North Europe, peculiar for CDD. In this thesis the interest lies in the fractional property of the temperature and the influence of the property on price of the weather derivatives. The prices of two types of contracts are simulated for 5 different H values each. The results of simulations turn out to be in two opposite ways. For a HDD contracts, the prices remain the same for varying H values. In case of a Froze contracts, the prices differ 74.37% from the $H = 0.5$ to $H = 0.9$. Such a dramatic behavior is however expected. The H values change only the consecrations of simulated pathes of the fBm, and the H values do not have any effect on expectation. The average temperature in January is around $-3°$. Even though the higher H values make the simulation more dispersed, the 18° is still unreachable. That is the reason why the H values have no effects on contrasts of HDD written on term January. That will be another story when the type of contract is Froze. The Froze index counts number of days, which the temperature is below zero. A dispersion of pathes of the simulation reduces the Froze index and makes price of a Call option based on the index lower. It is easier to understand if the dispersion is explained in the following way: it makes a warm day to be warmer and a cold day to be colder. The first case reduces the Froze index and the second case changes nothing for the Froze index. The conclusion is that the effect of the H values on price depends on the choice of the underlying indexes. For indexes which have the strike around the average temperature of the chosen term, the H values are significant factors for prices.

The thesis has succeeded in introducing a fractional Black-sholes market model for the temperature-based weather derivatives, both theoretically and through data analysis. The fBm is studied from the definition, the application in the finance market to the estimation of H values, and its simulation by statistic program R. The weather derivatives marked is studied from the market structure, the application in real-life to modeling temperature. The using the results both from the fBm and the weather derivatives market, the prices of derivatives are studied in two approaches: the analytical approach and the Monte Carlo.

7.2 Limits and future research

The thesis attempts to give an overall understanding of the fBm and the weather derivatives market as its typical example. However, by the limitation of the time, both of these two fields are too complex to cover. Many of the valuable subjects deserve more attentions.

The portfolio processes in the fractional Black-scholes market receive no attention. It is however an important part of the model. The calculation examples of the analytical price formulas may give more understanding of the price processes, and even the whole fractional Black-scholes model.

The connection of the continuous time fractional Ornstein-Uhlenbeck model to the discrete time AR(1) model can be studied in more details. The gap will be smaller if a consistent model is suggested from the beginning of the thesis. The difference from the theoretical part to the simulation part needs to be studied. The quasi-conditional expectation has no simple way of simulation. However the fast Fourier transform function in R, may prove an alternative way. The analytical price formulas are difficult to calculate, a R function can be made for that purpose. The difference of quasi-conditional expectation, quasi-martingale and the classical ones can be studied by using the same estimates for the analytical price formulas and the Monte Carlo approach. If the temperature data are selected from a city where the weather derivatives market is operative, the market price of the risk can also be calculated. If the market price of the risk is chosen, the analytical price and the Monte Carlo price can be compared to the realistic prices. The both pricing approaches can be much more improved.

More researches are needed for connecting the weather derivatives market to the potential participators. The car insurance example uses simple statistic and the results are not concluded. To reduce the basic risk, more sophisticated models are needed to investigate the relation between the temperature and car losses.

In this thesis, the meteorological forecast is not taken into account, and multivariable model for temperature is not considered.

The fractional Brownian motion and the weather derivatives marked are two relatively new fields. The author expects more cutting-edge results to be presented in the future.

Appendix A

Figure

Figure A.1: *Pay-off of call and put*

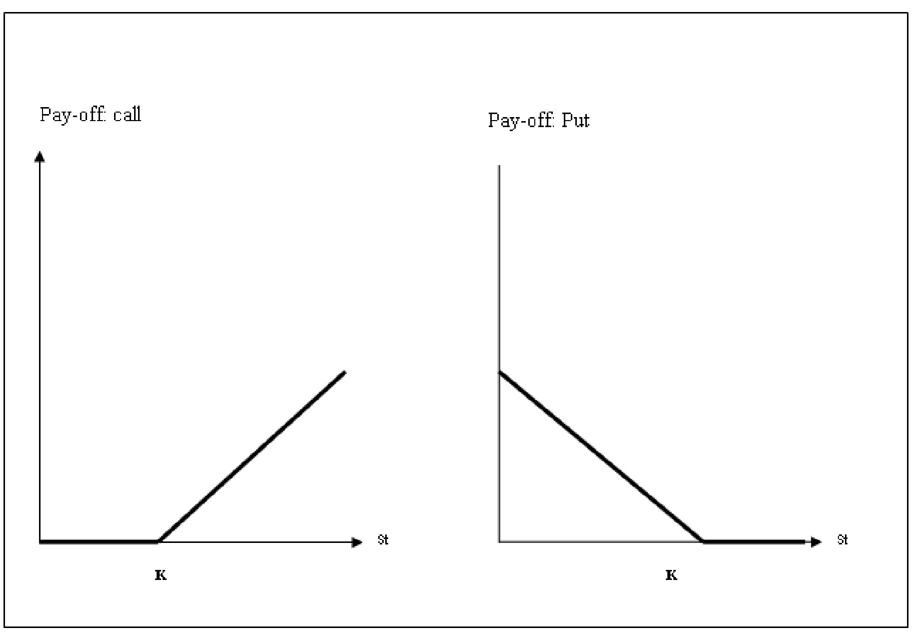


Figure A.2: *Pay-off of the example in table 3.1*

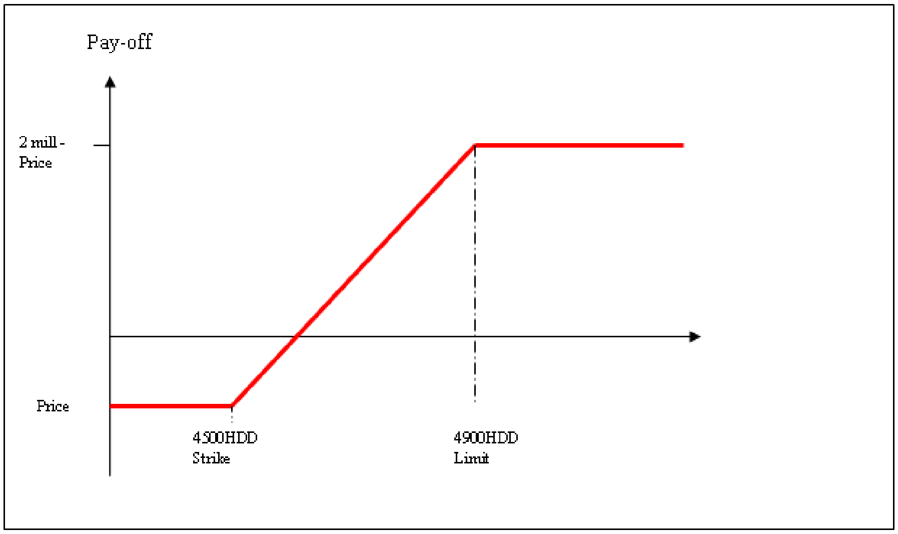


Figure A.3: *The growth of the numbers of policies*

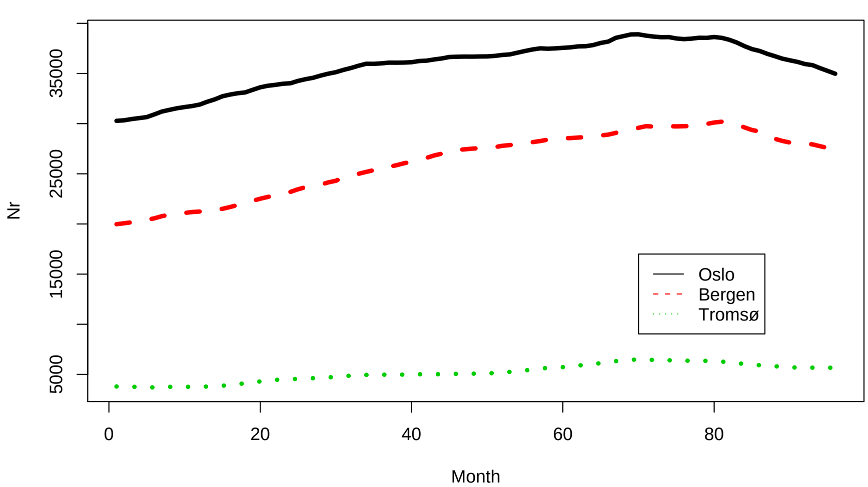


Figure A.4: *The number of losses, T in the first row and T-D-G in the second*

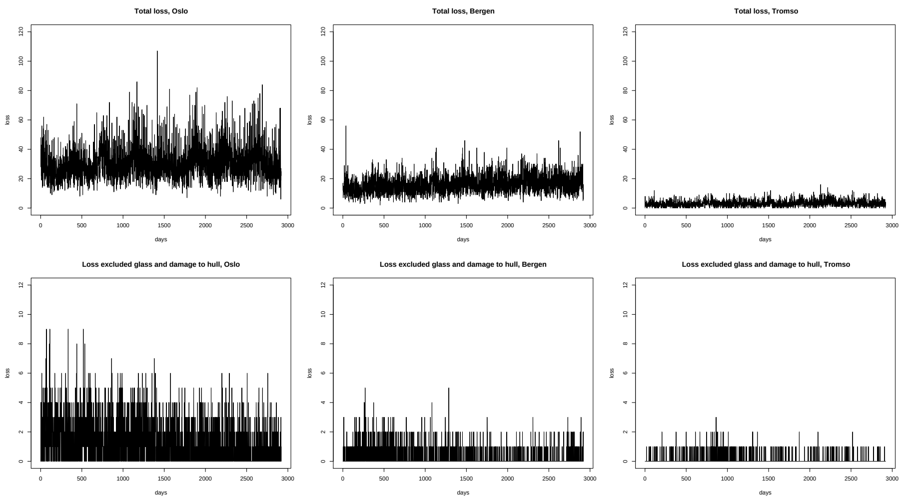


Figure A.5: *The 8 years average of the total losses, for Oslo, Bergen and Tromsø*

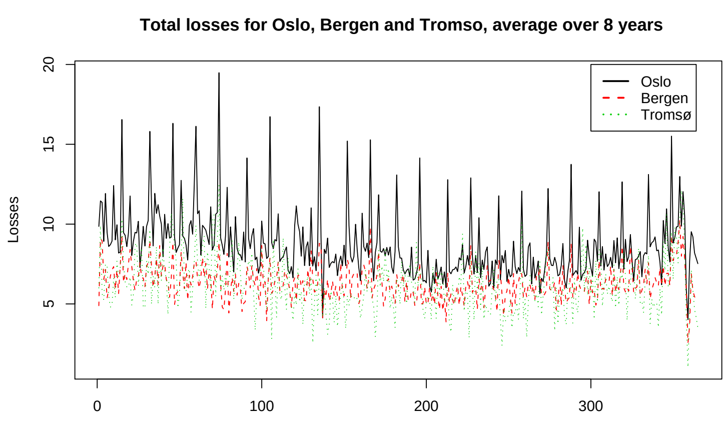


Figure A.6: *Average temperature, T and T-G-D, Oslo*

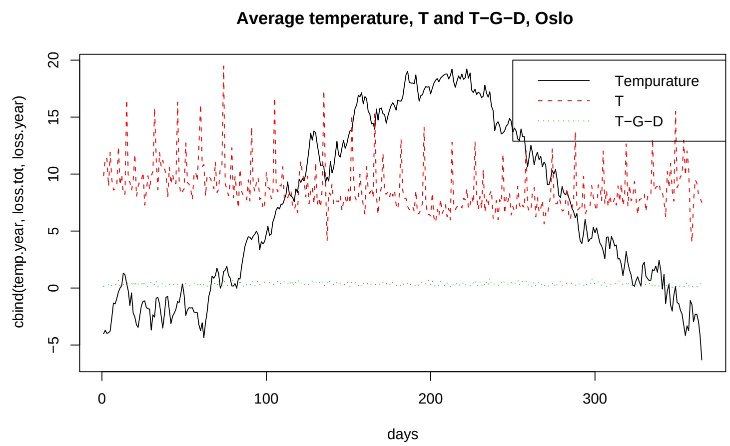


Figure A.7: *Boxplot, the 4 groups of temperature against the total losses*

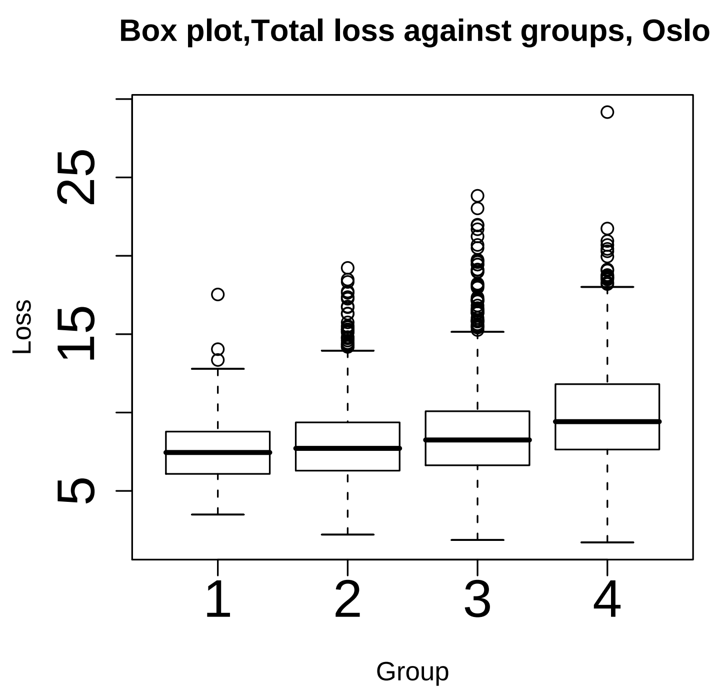


Figure A.8: *Histogram of daily average temperature from Oslo and Tromsø, together with the best fitted normal distributions*

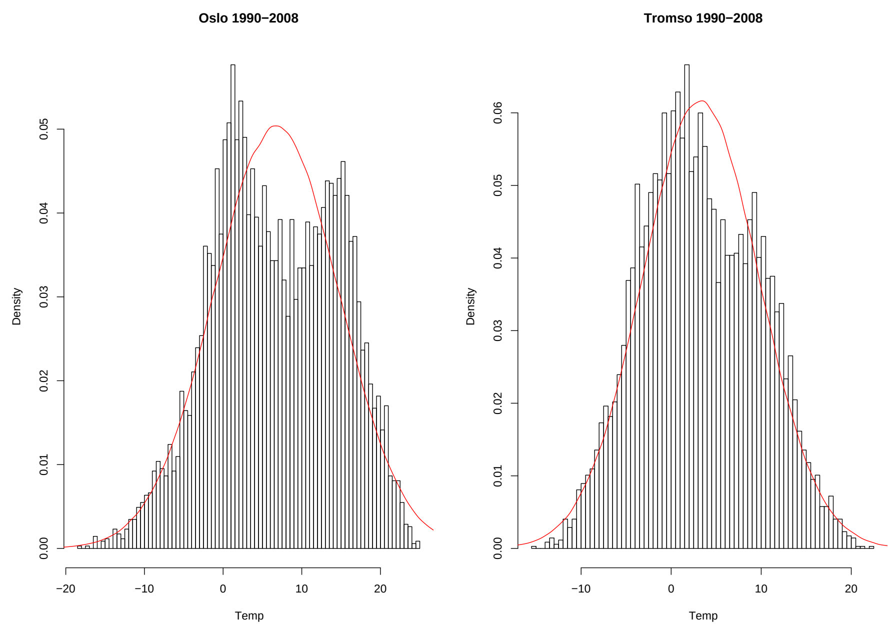


Figure A.9: *Daily average temperature from Oslo and Tromsø*

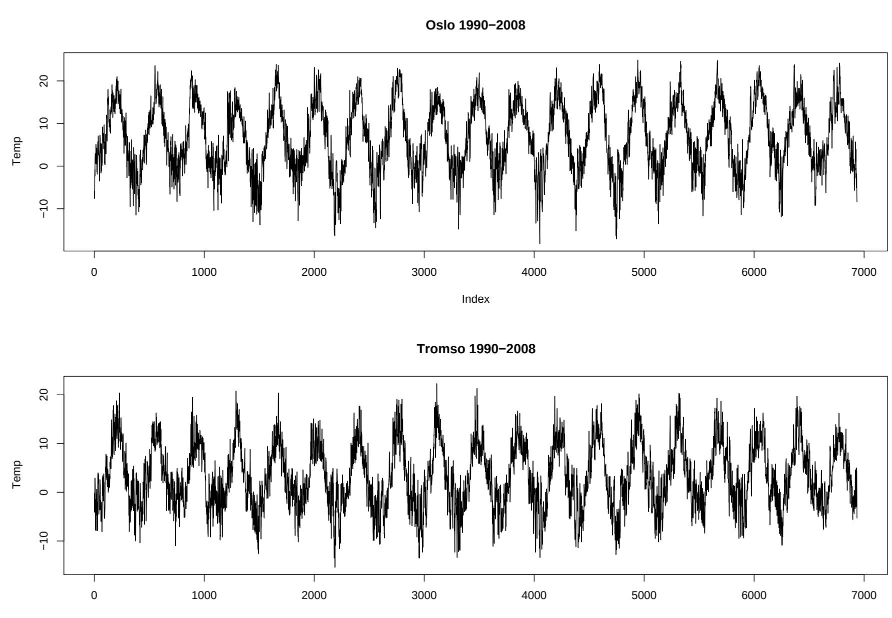


Figure A.10: *Detrended and deseasonalized daily average temperature from Oslo and Tromsø*

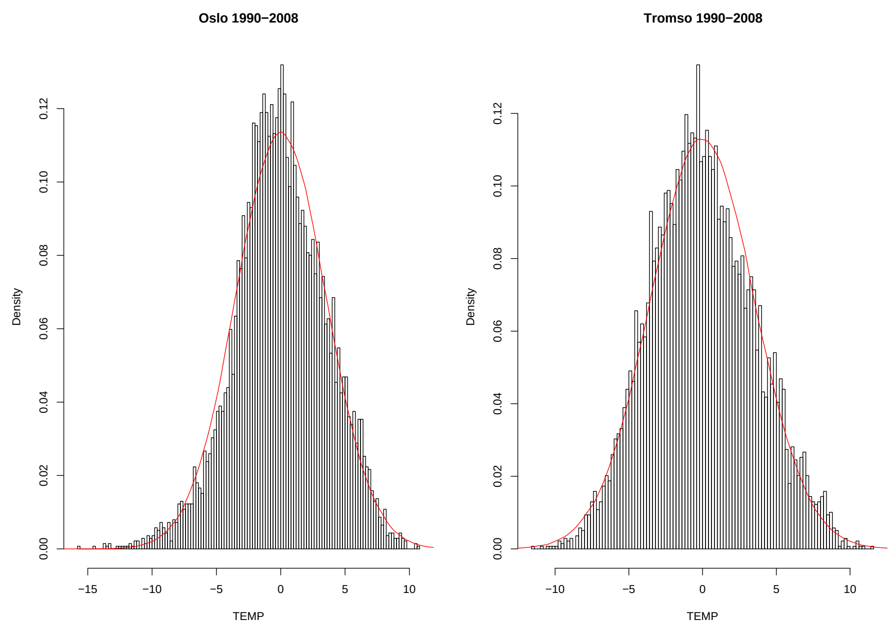


Figure A.11: *Daily average temperature, together with the estimated seasonal function and the detrended and deseasonalized temperature for Oslo*

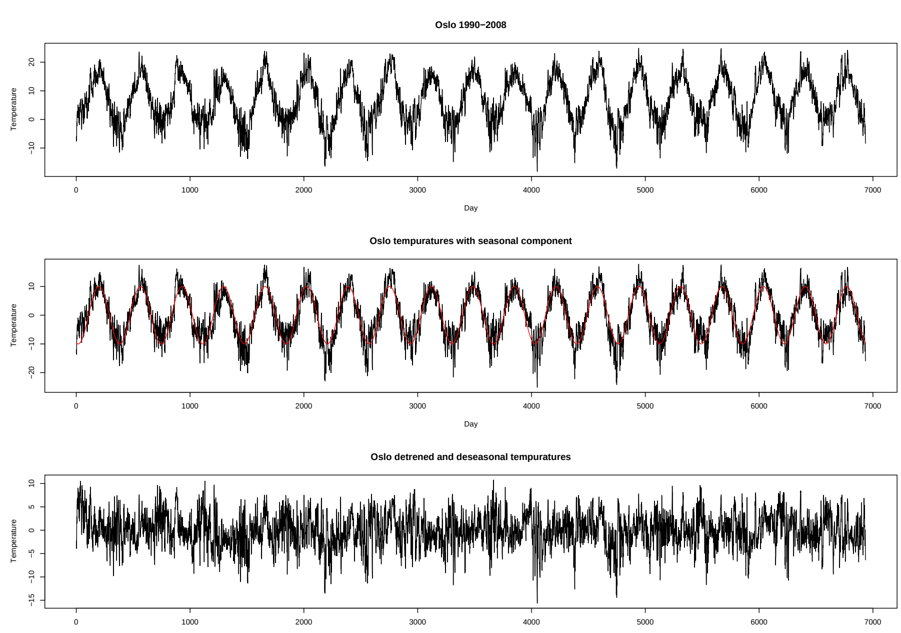


Figure A.12: *The residuals of Oslo and Tromsø*

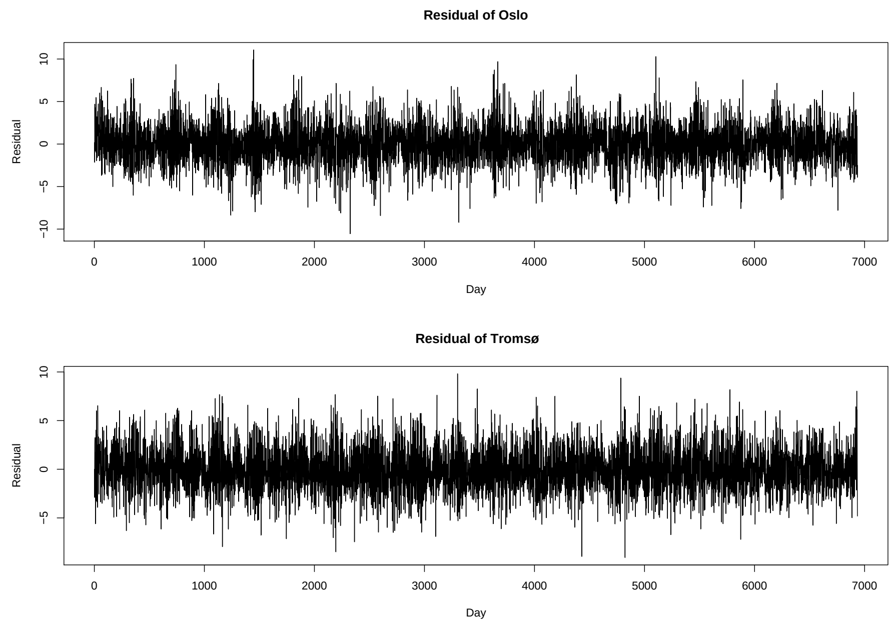


Figure A.13: *QQ normal plots for annual of the residual, together with QQ normal plot for 19 years, Røros*

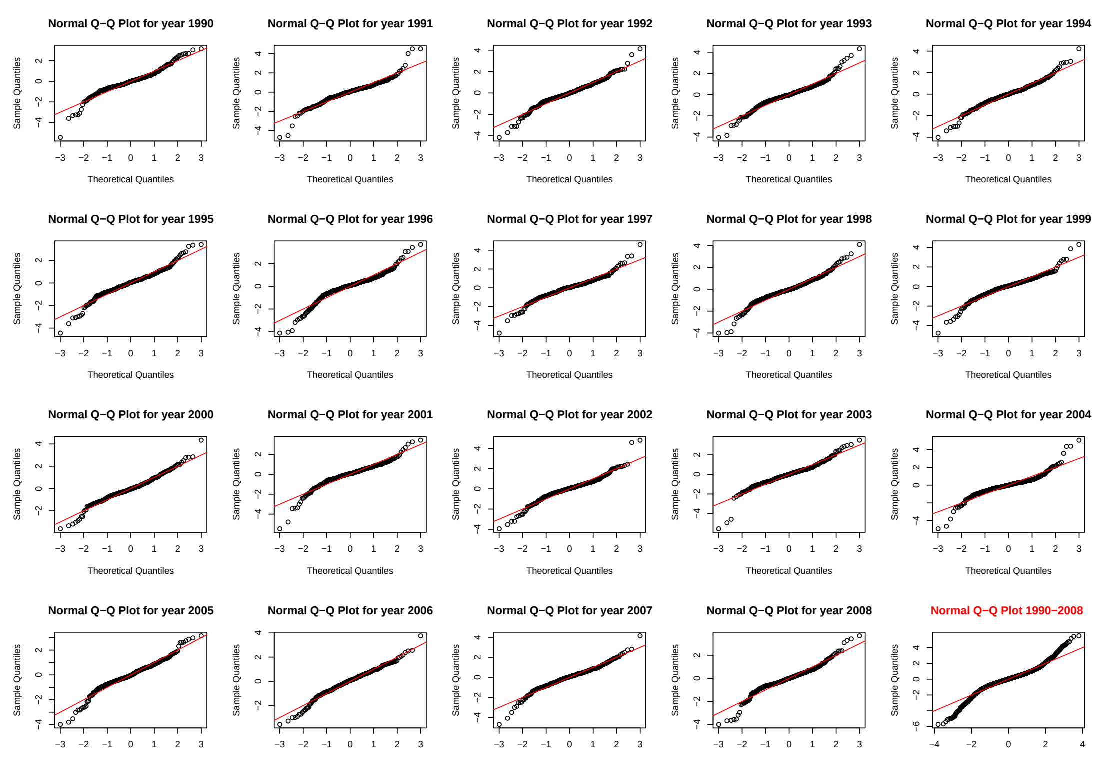


Figure A.14: *Empirical sigma for Bergen and Røros*

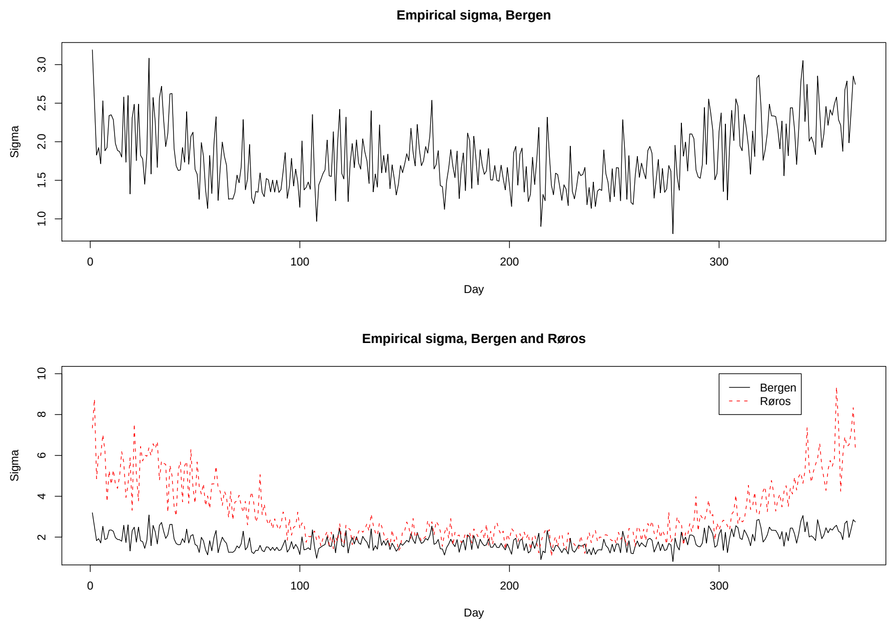


Figure A.15: *Empirical and estimate sigma for Bergen and Røros*

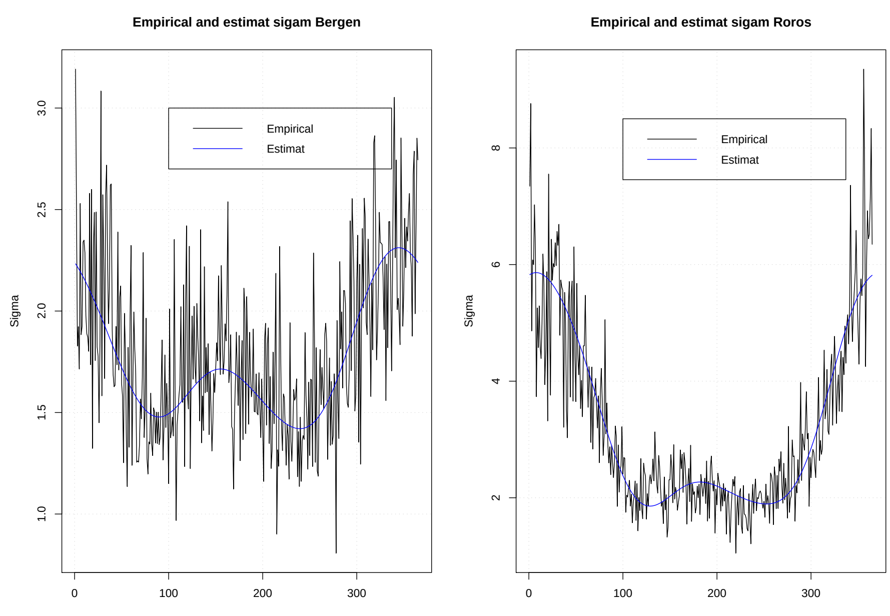


Figure A.16: *Plots of fBm, for H=0.1, H=0.5 and H=0.9, from top to bottom*

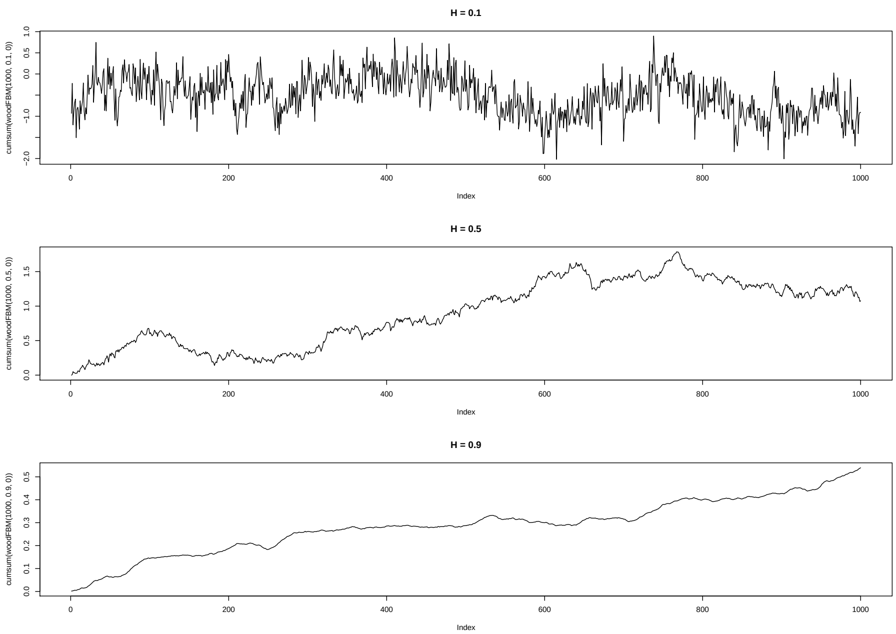


Figure A.17: *Biases and standard errors for the ST, RS and DFA method*

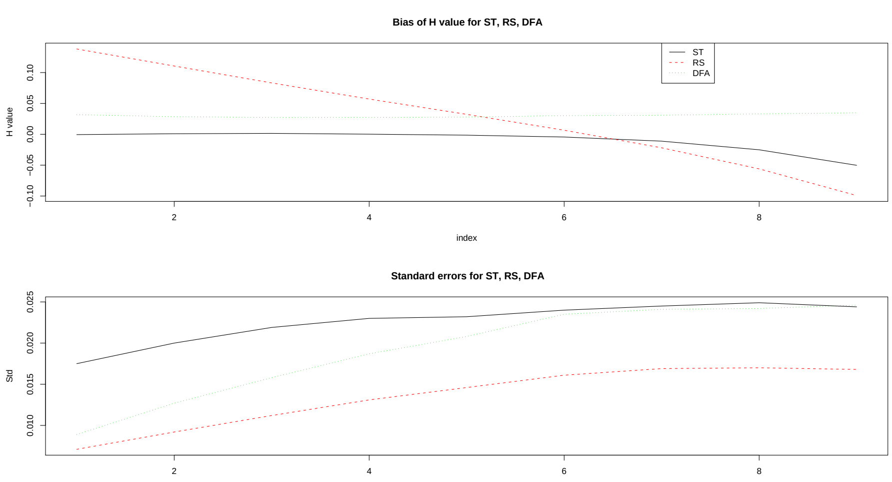


Figure A.18: *Plot of the ST, RS and DFA method, Oslo*

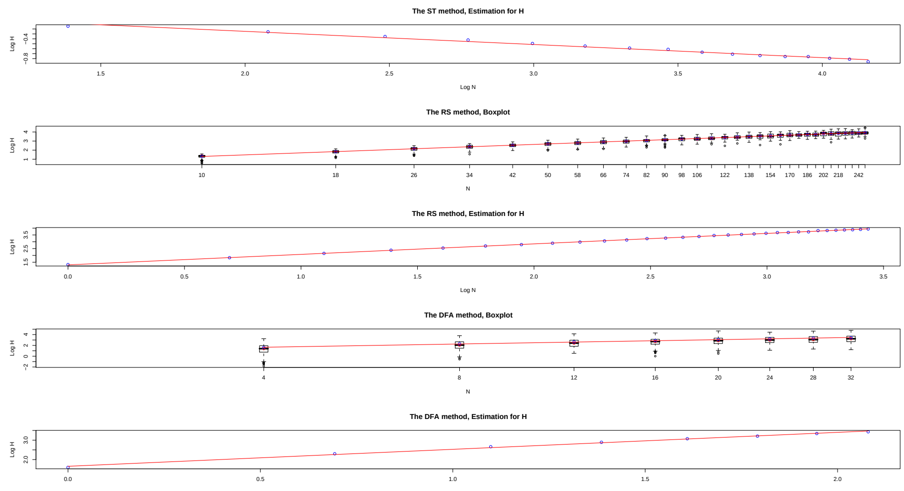


Figure A.19: *Autocorrelation for the 4 stage of temperature, Oslo*

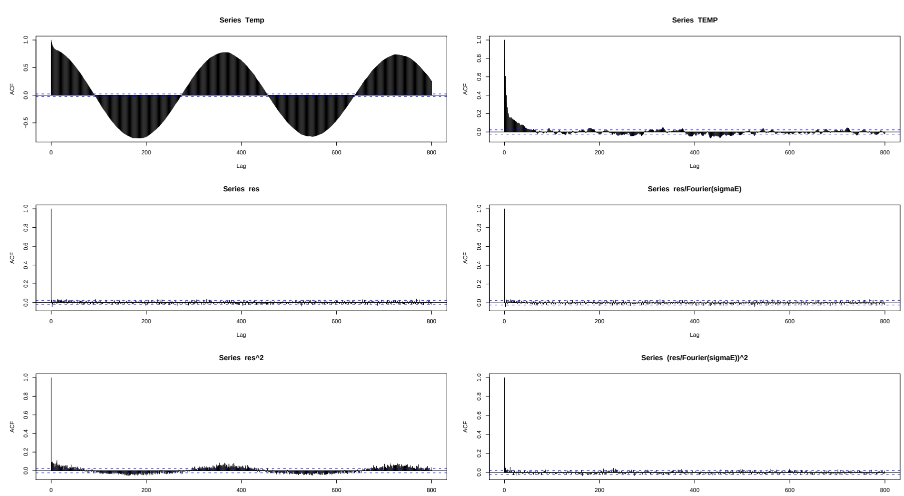
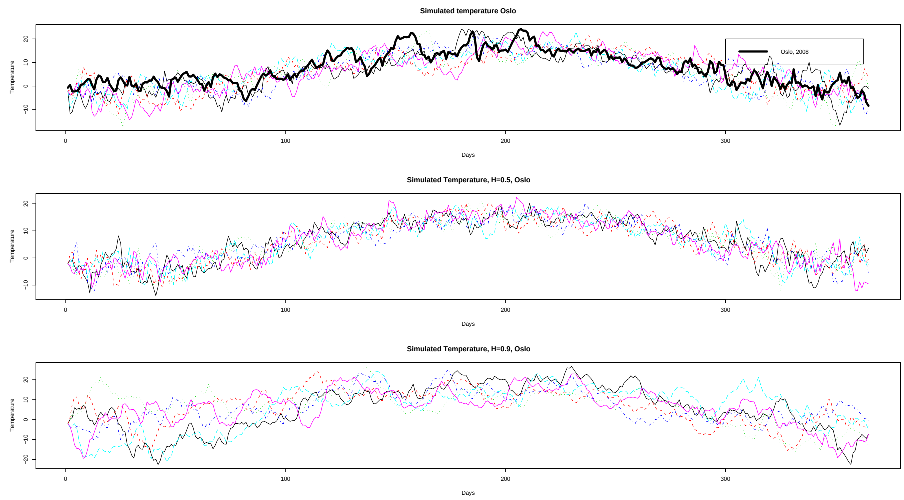


Figure A.20: *Simulated temperature with* $H = 0.598$,$H = 0.5$ *and* $H = 0.9$, *Oslo*



Appendix B

R scrips

B.1 A real-life example - Car insurance

```
## Car data
##-----------------------------------------------------------------------
##-----------------------------------------------------------------------

Car<-function(Loss,Temp, Policy){
##-----------------------------------------------------------------------
## Author :               Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Reference   :
## Description: Function check the relation between nr of icing days and cae damage
##
## Innput:                Loss     -        Nr of Loss
##                        Temp     -        Temperature
##                        Policy -     Nr of policy
##
## Output:
##
## Example:               Car(loss.Oslo,Temp.Oslo,policy.Oslo)
## -----------------------------------------------------------------------
        #-------------------------
        # Index for month
        #-------------------------
        mnd<-c(31,28,31,30,31,30,31,31,30,31,30,31)
        mnd<-rep(mnd,8)
        #-------------------------
        # Normaliraze from poliser
        #-------------------------
        #-------------------------
        # Trend
        #-------------------------
        poli.year<-rep(Policy,mnd)
        interval<-c(1:length(poli.year))
        p.trend<-lsfit(interval,poli.year)
        p0<-p.trend$coef[1]
        p1<-p.trend$coef[2]

        l.trend<-lsfit(interval,Loss)
        l0<-l.trend$coef[1]
        l1<-l.trend$coef[2]
        print(cbind(p0,p1,l0,l1))

        loss.norm<-(Loss/poli.year)*10000

        loss.year<-matrix(loss.norm,365,8)
        loss.year<-rowMeans(loss.year)

        temp.year<-matrix(Temp,365,8)
        temp.year<-rowMeans(temp.year)


        #matplot(cbind(loss.year,temp.year),type='l')
        null<-match(0,temp.year)
        abline(v=null, col='red')

        drop(loss.year)
}

Ice<-function(Loss,Temp, Policy){
##-----------------------------------------------------------------------
## Author :               Qingsheng Dong, UiO, 2009, Version 1.0 Beta
```

```
##
## Reference   :
## Description: Function check the relation between nr of icing days and cae damage
##
## Innput:              Loss     -        Nr of Loss
##                      Temp     -        Temperature
##                      Policy -     Nr of policy
##
## Output:
##
## Example:             Ice(loss.Oslo,Temp.Oslo,policy.Oslo)
## -----------------------------------------------------------------------------

        #------------------------
        # Catagolisert data,month
        #------------------------
        mnd<-c(31,28,31,30,31,30,31,31,30,31,30,31)
        mnd<-rep(mnd,8)
        index<-c(0,cumsum(mnd))
        #Group
        G<-0*(1:length(Temp))

        #------------------------
        # Normaliraze from poliser
        #------------------------
        poli.year<-rep(Policy,mnd)

        loss.norm<-(Loss/poli.year)*10000

        G[Temp<0]=4
        G[Temp==0]=4
        G[Temp>0&Temp<10]=3
        G[Temp==10]=3
        G[Temp>10&Temp<20]=2
        G[Temp==20]=2
        G[Temp>20]=1

        L.data  <- data.frame(list(group=as.factor(G)), Loss=c(loss.norm),poli=c(poli
             .year))
        boxplot(Loss~group,data=L.data,cex.axis=2,xlab='Group',ylab='Loss',main='Box
             plot,Total loss against groups, Oslo')
        fit.aov<-aov(Loss~group,data=L.data)
        fit.lm<-lm(Loss~group,data=L.data)
        print(summary(fit.aov))
        print(summary(fit.lm))

        #drop(Loss.lm)
}


        #----------------------------------
        # Read the data
        #----------------------------------
        policy<- read.delim2('policyKommune.txt',sep='')
        loss<- read.delim2('lossKommune.txt',sep='')

        tmpO<-read.delim2('Oslo19902008.txt',sep=";",dec=",",skip=17)
        tmpB<-read.delim2('Bergen19902008.txt',sep=";",dec=",",skip=17)
        tmpT<-read.delim2('Tromso19902008.txt',sep=";",dec=",",skip=17)

        #---------------------------------
        # Cut the data
        #---------------------------------
        # Nr
        policy.Oslo<-policy$Number[25:120]
        policy.Bergen<-policy$Number[145:240]
        policy.Tromso<-policy$Number[265:360]
        matplot(cbind(policy.Oslo,policy.Bergen,policy.Tromso),ylab='Nr',xlab='Month'
             ,type='l',lwd=2)


        # Loss
        #lossT<-loss$total-loss$glass-loss$damage
        lossT<-loss$total
        #lossT<-loss$damage
        loss.Oslo<-lossT[731:3650]
        loss.Bergen<-lossT[4381:7300]
        loss.Tromso<-lossT[8031:10950]


        # Zero the negativ nr
        loss.Oslo[loss.Oslo<0]=0
        loss.Bergen[loss.Bergen<0]=0
        loss.Tromso[loss.Tromso<0]=0


        # Temp
        a<-(365*11+1)
        b<-(365*19)
        Temp.Oslo<-tmpO$TAM[a:b]
```

```
    Temp.Bergen<-tmpB$TAM[a:b]
    Temp.Tromso<-tmpT$TAM[a:b]

    #---------------------------------------
    # Plot
    #---------------------------------------
    l.o<-Car(loss.Oslo,Temp.Oslo,policy.Oslo)
    l.b<-Car(loss.Bergen,Temp.Bergen,policy.Bergen)
    l.t<-Car(loss.Tromso,Temp.Tromso,policy.Tromso)
    matplot(cbind(l.o,l.b,l.t),type='l',xlab='days',ylab='Losses',main='Total losses for Oslo, Bergen and Tromso, average over 8 years')
```

B.2 Generator for fBm - The Wood-Chan's method

```
## Simulation Of fBm based on Wood-Chan's method
#-----------------------------------------------------------------------------
#-----------------------------------------------------------------------------


woodFBM<-function(N,H,Plot){
## ---------------------------------------------------------------------------
##
## Reference   :         1.Wood A. and Chan G. , Simulation of stationnary
##                       Gaussian processes, Journal of computational and
##                       graphical statistics, vol. 3, 1994
##                       2.Coeurjolly J.-F. , Simulation and identification
##                       of the fractional Brownian Motion: A biblographical
##                       and compapative study, J.Statist. Soft. vol. 5,
##                        page 1-53, 2000
##
## Description:          Function based on Wood-Chan's method and give a sample
##                       of fBm. Idea is using antocovariance matrix G to
##                       estimate fractional Gaussion noise(fGn) and culmulated
##                       sum of fGn given a fBm. The special part og the mehtod
##                       is using circulant matrix C to compute G
##
## Innput:               N - sample length
##                       H - Hurst coefficient
##
## Output:               Simulation of a fBm with length N and Hurst coefficient H
##
## Time:                 system.time(woodFBM(N=1000, H=0.8))= 0.08 sec
##
## --------------------------------------------------------------------------
    if(missing(Plot)) Plot<-1
#----------------------------------------
# Construction of C matrix
#----------------------------------------

    Cmatrix<-function(m,N,H){
    ##--------------------------------------------------------------------
    ## Innput:            m      - length og matrix C
    ##                    N      - Sample length
    ##                    H    - Hurst coefficient
    ## Output:            rowC   - First row of circulant matrix C which
    ##                             build by covariances of fBm
    ##--------------------------------------------------------------------
    j<-(0:(m-1))
    H2<-2*H
    rowc<-(abs((j-1)/N)^H2-2*(j/N)^H2+((j+1)/N)^H2)/2            #Covariances
    index<-c(0:(m/2-1),m/2,(m/2-1):1)                             #Index of row
    rowc<-rowc[index+1]
    drop(rowc)
    }

    #---------------------------------------
    # Find m = (2^?)>2(N-1)
    #---------------------------------------
    m<-2^(floor(log((N-1))/log(2)))

    #---------------------------------------
    # Find the minst positive definete C,
    # and dim<2^17
    #---------------------------------------
    repeat{
        m<-2*m
        ev<-Cmatrix(m,N,H)
        ev<-fft(ev,inverse = F)
        if((all(Re(ev)>0))|m>2^17) break
    }

    # Check m
    if(m>2^17){
```

```
                print('Method impossible')
                break
        }else{
        #------------------------------------------------
        # Simulation begin now
        #---------------------------------------------
                re<-rnorm(m/2+1)
                im<-rnorm(m/2+1)
                #re[1]<- sqrt(2)*re[1]
                #re[(m/2+1)]<-sqrt(2)*re[(m/2+1)]
                #im[1]<-0
                #im[(m/2+1)]<-0
                re<-c(re[1:(m/2+1)],re[(m/2):2])
                imc<- -im
                im<-c(im[1:(m/2+1)],imc[(m/2:2)])
                W<-complex(real=re,imaginary=im)/sqrt(2)
                W[1]=re[1]
                W[m/2+1]=im[1]

        #------------------------------------------------
        # fGn
        #------------------------------------------------
                W<- sqrt(ev)*W
                fGn<-fft(W)
                fGn<-(1/(sqrt(m)))*fGn
                fGn<-Re(fGn[c(1:N)])

        #------------------------------------------------
        # fBm
        #------------------------------------------------
                fBm<- fGn

        #------------------------------------------------
        # Plot
        #------------------------------------------------
                if(Plot==1){
                        plot(fBm,type='l')
                }
                drop(fBm)
        }
}

tt<-woodFBM(100,0.5)
```

B.3 Estimator for H values

ST method

```
## Estimate H coefficient basert on ST method
##------------------------------------------------------------------------------
##------------------------------------------------------------------------------

ST<-function(TEMP,MC){
##------------------------------------------------------------------------------
## Author :              Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Reference  :          Dorje C Brody et al. Dynamical pricing of weather
##                       derivatives, Quantitative finance, Volume 2, 2002, 189-198
##
## Description: Function based on ST method, introduced by Syroka and Toumi, to
##                       estimate H coefficient
##
## Innput:              TEMP    - Data, fBm Sample, no missing value
##                      MC      - If 1, Use MC and no plot
##
## Output:              H       - Hurst coefficient
##
## Example:             ST(woodFBM(N=1000, H=0.8))
## Time:                10000 x system.time(ST(woodFBM(N=1000, H=0.6)))=109 sec
## Quality:             mean=0.590935, var=0.02770983,  for MC 10000 times
##
## -----------------------------------------------------------------------------
        if(missing(MC)) MC<-0

#------------------------------------------------
# Estimate H for time interval of T days
#------------------------------------------------

        FindH<-function(T,Temp){
        ##----------------------------------------------------------------------
        ## Innput:        T       - Length of interval
        ##                Temp    - Data
        ## Output:        H       - Hurst coefficient
        ##----------------------------------------------------------------------
```

```
                    #---------------------------------------------------------------------
                    # Divide data into N bins, and find starting index for each bins
                    #---------------------------------------------------------------------
                    N<-round(length(Temp)/T)-1
                    TempT<-Temp[(length(Temp)-T*N):length(Temp)]
                    index<- seq(1,(T*(N+1)),by=T)

                    M<-0*c(1:N)

                    ## Put data in M
                    for(ii in 1:N){
                            M[ii]=mean(TempT[index[ii]:(index[ii+1]-1)])
                    }

                    MTemp<-mean(TempT)                                      #Total average
                    STemp<-sd(TempT)                                              #Total sd
                    sigma<-sqrt((1/(length(M)))*sum(((M-MTemp)/STemp)^2))     #Sigma(T)

                    drop(sigma)
            }

            #---------------------------------------------------------------------
            # Try different values of T
            #---------------------------------------------------------------------
            Nr<-seq(4,64,by=4)
            N<-round(length(TEMP)/Nr)-1

            HH<-0*c(1:length(Nr))
            for(ii in 1:length(Nr)){
                    HH[ii]<- FindH(Nr[ii],TEMP)
            }

            lm.ST<-lm(log(HH)~log(Nr))
            coe<-lm.ST$coefficient

            if(MC!=1){
                    plot(y=log(HH),x=log(Nr),xlab='Log N',ylab='Log H',main='The ST 
                        method, Estimation for H',col='blue',type='p')
                    lines(y=coe[1]+coe[2]*log(Nr),x=log(Nr),type='l',col='red')
            }
            drop(1+coe[2])
}


##---------------------------------------------------------------------
## MC for method, sim = 10000
##---------------------------------------------------------------------
MC<-0

if(MC==1){
            sim<-10000
            aveHH<-0*(1:sim)
            for(ss in 1:sim){
                    aveHH[ss]<-ST(woodFBM(1000,0.8,2),1)
            }
            print(mean(aveHH))
            print(sd(aveHH))
}
```

RS method

```
## RS method
## Estimate H coefficient basert on RS method
##---------------------------------------------------------------------
##---------------------------------------------------------------------

RS<-function(TEMP,MC){
##---------------------------------------------------------------------
## Author :       Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Reference   : FOTINI PALLIKARI AND EMIL BOLLER,
##                A Rescaled Range Analysis of Random Events, Journal of
##                Scientific Exploration, Vol. 13, No. 1, pp. 25-40, 1999
##
## Description: Function based on RS method, introduced by Hurst,
##        to estimate H coefficient
##
## Innput:        TEMP     - Data, fBm Sample, no missing value
##                         MC                 - If 1, Use MC and no plot
##
## Output:                 H                  - Hurst coefficient
##
## Example:                                   S(woodFBM(N=1000, H=0.8))
## Time:                   10000 x system.time(RS(woodFBM(N=1000, H=0.6)))= 250 sec
## Quality:                mean=0.6066, var=0.016,  for MC 10000 times
```

```
##
## ---------------------------------------------------------------------------------------
        if(missing(MC)) MC<-0

        #----------------------------------------------
        # Estimate RS for each bins
        #----------------------------------------------
        RSN<-function(XN){
                N<-length(XN)
                m<-mean(XN)
                XtN<-0*c(1:N)
                XtN<-cumsum(XN-m)
                RN<-max(XtN)-min(XtN)
                SN<-sqrt((1/N)*sum((XN-m)^2))
                rsn<-RN/SN
        }

        #----------------------------------------------
        # Length of bins
        #----------------------------------------------
        Lval<-seq(10,256,by=8)
        Nval<-trunc(rep(length(TEMP),length(Lval))/Lval)

        #----------------------------------------------
        # Find RS for different length of bins,
        #----------------------------------------------
        Box.data<-NULL
        LRS<-0*(1:length(Lval))
        LRSM<-0*(1:length(Lval))
        for(ii in 1:length(Lval)){
                TempT<-TEMP[(length(TEMP)-Lval[ii]*Nval[ii]):length(TEMP)]

                #----------------------------------------
                # Calculete E(R/s)
                #----------------------------------------
                index<- seq(1,(Lval[ii]*Nval[ii]+1),by=Lval[ii])
                rs<-c(1:Nval[ii])
                for(jj in 1:Nval[ii]){
                        XN<-TEMP[index[jj]:(index[jj+1]-1)]
                        rs[jj]<-RSN(XN)
                }
                LRS[ii]<-log(mean(rs))
                LRSM[ii]<-median(log(rs))
                Box.data  <- c(Box.data,log(rs))
        }

        #----------------------------------------------
        # Linear regression
        #----------------------------------------------
        Box.group <- as.factor(rep(Lval, times = Nval))
        Box.data  <- data.frame(list(group = Box.group, Box = Box.data))
        xval<-c(1:length(Nval))
        lm.RS<-lm(LRS~log(xval))
        coe<-lm.RS$coefficient

        #----------------------------------------------
        # Boxplot,plot, print
        #----------------------------------------------
        if(MC==0){
                par(mfrow=c(2,1))
                boxplot(Box ~ group, data = Box.data,log='x',boxwex = 0.03,xlab='N',
                     ylab='Log H',main='The RS method, Boxplot')
                lines(y=LRSM,x=xval,col='red',type='p')
                lines(y=LRS,x=xval,col='blue',type='p')
                lines(y=coe[1]+coe[2]*log(xval),x=xval,type='l',col='red')

                plot(y=LRS,x=log(xval),col='blue',type='p',xlab='Log N',ylab='Log H',
                     main='The RS method, Estimation for H')
                lines(y=coe[1]+coe[2]*log(xval),x=log(xval),type='l',col='red')

                print('Coefficient')
                print(coe)
       }

        drop(coe[2])
}



##------------------------------------------------------------------------------
## MC for method, sim = 10000
##------------------------------------------------------------------------------
MC<-0

if(MC==1){
        sim<-10000
        aveHH<-0*(1:sim)
        for(ss in 1:sim){
                aveHH[ss]<-peraggST(cumsum(woodFBM(1000,0.5,3)))
```

```
        }
        print(mean(aveHH))
        print(sd(aveHH))
}
```

DFA method

```
## DFA method
## Estimate H coefficient basert on RS method
##--------------------------------------------------------------------------------
##--------------------------------------------------------------------------------

DFA<-function(TEMP,MC){
##--------------------------------------------------------------------------------
## Author :       Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Reference  :   Blocks adjustment-reduction of bias and variance of detrended
## fluctuation analysis using Monte Carlo simulation,
## Michalski, Sebastian, Physica A 387, pages 217-242. 2007
##
## Description: Function based on RS method, introduced by Hurst, to
##                      estimate H coefficient
##
## Innput:              TEMP    - Data, fBm Sample, no missing value
##                      MC      - If 1, Use MC and no plot
##
## Output:              H       - Hurst coefficient
##
## Example:             RS(woodFBM(N=1000, H=0.8))
## Time:                10000 x system.time(DFA(woodFBM(N=1000, H=0.6)))= 450 sec
## Quality:             mean=0.6302, var=0-0235,  for MC 10000 times
##
## --------------------------------------------------------------------------------
        if(missing(MC)) MC<-0

        #----------------------------------------------
        # Estimate Fn for each bins
        #----------------------------------------------
        Fn<-function(XN){
                Y<-cumsum(XN)
                interval<-c(1:length(XN))
                trend<-lsfit(interval,XN)
                b0<-trend$coef[1]
                b1<-trend$coef[2]
                Z<-Y-(b0+interval*b1)

                fn<-sqrt((1/length(XN))*sum(Z^2))
        }

        #----------------------------------------------
        # Length of bins
        #----------------------------------------------
        Lval<-seq(4,32,by=4)
        Nval<-trunc(rep(length(TEMP),length(Lval))/Lval)

        #----------------------------------------------
        # Find Fn for different length of bins,
        #----------------------------------------------
        Box.data<-NULL
        LFN<-0*(1:length(Lval))
        LFNM<-0*(1:length(Lval))
        for(ii in 1:length(Lval)){
                TempT<-TEMP[(length(TEMP)-Lval[ii]*Nval[ii]):length(TEMP)]

                #----------------------------------------
                # Calculete E(Fn)
                #----------------------------------------
                index<- seq(1,(Lval[ii]*Nval[ii]+1),by=Lval[ii])
                F<-c(1:Nval[ii])
                for(jj in 1:Nval[ii]){
                        XN<-TEMP[index[jj]:(index[jj+1]-1)]
                        F[jj]<-Fn(XN)
                }
                LFN[ii]<-log(mean(F))
                LFNM[ii]<-median(log(F))
                Box.data  <- c(Box.data,log(F))
        }

        #----------------------------------------------
        # Linear regression
        #----------------------------------------------
        Box.group <- as.factor(rep(Lval, times = Nval))
        Box.data  <- data.frame(list(group = Box.group, Box = Box.data))
        xval<-c(1:length(Nval))
        lm.FN<-lm(LFN~log(xval))
```

```
        coe<-lm.FN$coefficient

        #----------------------------------------------
        # Boxplot, plot,  print
        #----------------------------------------------
        if(MC==0){
                par(mfrow=c(2,1))
                boxplot(Box ~ group,  data = Box.data,log='x',boxwex = 0.03,xlab='N',
                    ylab='Log␣H',main='The␣DFA␣method,␣Boxplot')
                lines(y=LFNM,x=xval,col='red',type='p')
                lines(y=LFN,x=xval,col='blue',type='p')
                lines(y=coe[1]+coe[2]*log(xval),x=xval,type='l',col='red')

                plot(y=LFN,x=log(xval),col='blue',type='p',xlab='Log␣N',ylab='Log␣H',
                    main='The␣DFA␣method,␣Estimation␣for␣H')
                lines(y=coe[1]+coe[2]*log(xval),x=log(xval),type='l',col='red')

                print('Coefficient')
                print(coe)
        }

        drop(coe[2])
}

##------------------------------------------------------------------
## MC for method,  sim = 10000
##------------------------------------------------------------------
MC<-0

if(MC==1){
        sim<-10000
        aveHH<-0*(1:sim)
        for(ss in 1:sim){
                aveHH[ss]<-peraggST(cumsum(woodFBM(1000,0.5,3)))
        }
        print(mean(aveHH))
        print(sd(aveHH))
}
```

B.4 Estimation and simulation for temperature

Detrend and deseasonal

```
## Deseasonal data
##-------------------------------------------------------------------
##-------------------------------------------------------------------

Deseason<-function(City,Plot){
##-----------------------------------------------------------------
## Author :      Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Reference  :    Fred Espen Benth and Jurate Saltyte Benth,Stochastic modelling
##               of temperature variations with a view towards weather derivatives
##
## Description: Function remove seasonal cycle and detrended data samples
##
## Innput:               City    - City of data samples
##                       Plot    - If 1, plot data samples with seasonal cycle
##
## Output:               TEMP    - Detrended data sample without seasonal cycle
##                       Coef    - Coefficient for detrend and deseasonal
##                       c(b0,b1,a0,a1,t0)
##
## Example:              Deseason('Oslo',1)
## -----------------------------------------------------------------
        if(missing(Plot)) Plot<-1

        #Title='C:\\Documents and Settings\\u32018\\My Documents\\Master\\R\\'
        #Title=paste(Title,City,'19902008.txt',sep='')
        Title=paste(City,'19902008.txt',sep='')

        library(MASS)
        library(fBasics)
        library(tseries)

        #------------------------------------
        # Read inn data
        #------------------------------------
        colNames <- c("TAM","TAN","TAX")
        tmp <- read.delim2(Title,sep=";",dec=",",skip=17)

```

```
    #------------------------------------------------
    # Mean temperature
    #------------------------------------------------
    Temp<-tmp$TAM

  print(round(cbind(mean(Temp),max(Temp),min(Temp),sd(Temp),skewness(Temp),
      kurtosis(Temp)),2))
    # Normality test
  print(jarque.bera.test(Temp))

    #------------------------------------------------
    # Detrende, remove
    # least mean squared linear trend
    #------------------------------------------------
    interval<-c(1:length(Temp))
    trend<-lsfit(interval,Temp)
    b0<-trend$coef[1]
    b1<-trend$coef[2]
    print(cbind(b0,b1))
    TempT<-Temp-(b0+interval*b1)

    #------------------------------------------------
    # Seasonality
    #------------------------------------------------
    Season<-function(pare){
        a0<-pare[1]
        a1<-pare[2]
        t0<-pare[3]
        a0+a1*cos(((2*pi)/365)*(interval-t0))
    }

    Lik<-function(pare){
        sum(abs(TempT-Season(pare))^2,na.rm=TRUE)
    }

    #------------------------------------------------
    # Estimate for seasonal effekt
    #------------------------------------------------
    pareCS<-nlminb(objective = Lik,start=c(2,-3,10))
    LS<-pareCS$par
    print(round(LS,6))
    TempS<-TempT-Season(LS)

    #------------------------------------------------
    # Plot original, detrended, deseasonal
    # data samples and seasonal cycle
    #------------------------------------------------
    if(Plot==1){
        par(mfrow=c(3,1))

        pn<-paste(City,'1990-2008')
        plot(Temp,type='l',main=pn,xlab='Day',ylab='Temperature')

        pn<-paste(City,'tempuratures with seasonal component')
        plot(TempT,type='l',,main=pn,xlab='Day',ylab='Temperature')
        lines(Season(LS),type='l',col='red')

        pn<-paste(City,'detrened and deseasonal tempuratures')
        plot(TempS,main=pn,type='l',xlab='Day',ylab='Temperature')
    }

    #--------------------------------------------------
    # Output
    #--------------------------------------------------
    TEMP<-TempS[!is.na(TempS)]
    Coef<-c(b0,b1,LS)
    Out <- new.env()
    Out$TEMP <- TEMP
    Out$Coef <- Coef
    as.list(Out)
}

#Output<-Deseason('Oslo',1)
#TEMP<-Output$TEMP
#Coef<-Output$Coef
#print(Coef)
```

Simulation of temperature

```
## Simulation of Temperature
##------------------------------------------------------------------------------------------
##------------------------------------------------------------------------------------------

```

```
SimTEMP<-function(City,N0,N,Sim,Plot){
##-------------------------------------------------------------------------------------------
## Author :               Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Reference  :    1.Dorje C Brody et al. Dynamical pricing of weather derivatives,
##                 Quantitative finance, Volume 2, 2002, 189-198
##                 2.Benth, Fred Espen. On arbitrage-free pricing of weather derivatives
##                 based on fractional brownian motion
##
## Description: Function simulate temperature drived by fBm and follow
##                        Ornstein-uhlenbeck process
##
## Innput:                City      - City of data samples
##                        N0        - Starting days
##                        N         - Days of simulastion
##                        Plot      - If 1, plot data samples with seasonal cycle
##
## Output:                X.TEMP - Simulation of temperature
##
## Example:               SimTEMP('Oslo',1,10,1000,1)
##
## ------------------------------------------------------------------------------------------
        if(missing(N0)) N0<-1
        if(missing(N)) N<-365
        if(missing(Sim)) Sim<-500
        if(missing(Plot)) Plot<-1

        #-----------------------------------------------
        # Estimate for trend and season
        #-----------------------------------------------
        data<-Deseason(City)
        trend.coef<-data$Coef[1:2]
        season.coef<-data$Coef[3:5]
        TEMP<-data$TEMP
        print(cbind(trend.coef))
        print(cbind(season.coef))

        #-----------------------------------------------
        # Estimate for kappa
        #-----------------------------------------------
        L<-length(TEMP)
        AR1<-ar(TEMP, order.max=1)
        alpha<-AR1$ar
        kappa<--log(alpha)
        res<-AR1$resid
        res[1]=0
        print('AR1')
        print(AR1)

        #-----------------------------------------------
        # Estimate sigma_t
        #-----------------------------------------------
        col<-length(res)/365
        mSigma<-matrix(res,365,col)
        sigma<-0*(1:365)
        for(ii in 1:365){
                sigma[ii]<-sd(mSigma[ii,])
        }

        interval<-c(1:length(sigma))
        Fourier<-function(pare){
                c<-pare[1]
                c1<-pare[2]
                c2<-pare[3]
                c3<-pare[4]
                d1<-pare[5]
                d2<-pare[6]
                d3<-pare[7]
                c+c1*sin(((2*pi)/365)*interval)+c2*sin(((4*pi)/365)*interval)+
                c3*sin(((8*pi)/365)*interval)+
                d1*cos(((2*pi)/365)*interval)+d2*cos(((4*pi)/365)*interval)+
                d3*cos(((8*pi)/365)*interval)
        }

        Lik<-function(pare){
                sum(abs(sigma-Fourier(pare)))
        }

        #------------------------------------
        # Estimate for seasonal effekt
        #------------------------------------
        pareSF<-nlminb(objective = Lik, start=c(0,0,0,0,0,0,0))
        sigmaE<-pareSF$par

        #------------------------------------
        # Similation of temperature
        #------------------------------------
        Season<-function(pare){
                a0<-pare[1]
                a1<-pare[2]
```

```
                t0<-pare[3]
                a0+a1*cos(((2*pi)/365)*(interval-t0))
        }

        Start<-length(TEMP)-(365-N0)
        X.TEMP<-array(0,c(N,Sim))
        X.de<-array(0,c(N,Sim))
        X.de[1,]<-array(TEMP[Start],Sim)                #Starting point
        X.res<-array(0,c(N,Sim))
        X.error<-array(0,c(N,Sim))
        H<-ST((res/Fourier(sigmaE)),0)
        print(H)
  for(ii in 1:Sim){
                fBm<-woodFBM(N,H,0)
                X.error[,ii]<-(fBm-mean(fBm))/sd(fBm)   #Standarize
        }
        print(cbind(N0,N))
        interval<-(N0:(N+N0-1))
        X.sigma<-Fourier(sigmaE)
        X.res<-X.sigma*X.error

        for(ii in 2:N){
                X.de[ii,]<-alpha*X.de[(ii-1),]+X.res[ii-1,]
        }


        X.TEMP<-X.de+(trend.coef[1]+interval*trend.coef[2])+Season(season.coef)

        if(Plot==1){
                par(mfrow=c(3,1))

                plot(sigma,type='l')
                interval<-c(1:length(sigma))
                lines(Fourier(sigmaE),type='l',col='blue')

                plot(TEMP[1:N],type='l')
                lines(X.de[,3],type='l',col='red')
                #lines(X.de[,4],type='l',col='blue')

                matplot(X.TEMP[,5:10],type='l')
        }

        drop(X.TEMP)
}
```

B.5 Pricing, HDD, CDD, CAT and OTC for put and call

```
Price<-function(WD,N0,N,City,Sim,K,Plot){
##-----------------------------------------------------------------------------------
## Author :               Qingsheng Dong, UiO, 2009, Version 1.0 Beta
##
## Description:           Function give a price based on Monte Carlo approch.
##
## Innput:                WD     - Types of Weather Derivatives. HDD, CDD or CAT
##                        N0     - Starting point
##                        N      - Term i days, suggest 1 or 3 mnd
##                        City   - City of data samples
##                        Sim    - Nr of the simulation
##                        K      - Strike
##                        Plot   - If 1, plot data samples with seasonal cycle
##
## Output:                Index - Simulert index
##                        price.call - call price
##                        price.put  - put price
##
## Example:               Price('HDD',1,31,'Oslo',100000,620,1)
## ----------------------------------------------------------------------------------
        if(missing(WD)) WD<-'HDD'
        if(missing(N0)) N<-1
        if(missing(N)) N<-30
        if(missing(City)) City<-'Oslo'
        if(missing(Plot)) Plot<-1


        HDD<-function(TEMP){
                pmax(18-TEMP,0)
        }
        CDD<-function(TEMP){
                pmax(TEMP-18,0)
        }
        CALL<-function(INDEX,K){
                pmax(INDEX-K,0)
```

```
        }
        PUT<-function(INDEX,K){
                pmax(K-INDEX,0)
        }
        OTC<-function(TEMP){
                otc<-array(0,c(N,Sim))
                otc[(TEMP<0)]=1
                drop(otc)
        }
        #----------------------------------------------------
        # index of the year
        # 0   31   59   90 120 151 181 212 243 273 304 334 365
        #----------------------------------------------------
        r<-0.008
        hdd<-0*(1:N)
        price.call<-0
        price.put<-0
        if(WD=='HDD'){
                X.TEMP<-SimTEMP(City,N0,N,Sim,1)
                hdd<-colSums(HDD(X.TEMP))
                Index<-mean(hdd)
                print('Index')
                print(Index)

                price.call<-exp(-r*(N))*mean(CALL(hdd,K))
                price.put<-exp(-r*(N))*mean(PUT(hdd,K))
                pric.call.sd<-sd(CALL(hdd,K))
                pric.put.sd<-sd(PUT(hdd,K))
                print(cbind(price.call,price.put,pric.call.sd,pric.put.sd))
        }

        cdd<-0*(1:N)
        if(WD=='CDD'){
                X.TEMP<-SimTEMP(City,N0,N,Sim,1)
                cdd<-colSums(CDD(X.TEMP))
                Index<-mean(cdd)
                print('Index')
                print(Index)

                price.call<-exp(-r*(N))*mean(CALL(cdd,K))
                price.put<-exp(-r*(N))*mean(PUT(cdd,K))
                pric.call.sd<-sd(CALL(cdd,K))
                pric.put.sd<-sd(PUT(cdd,K))
                print(cbind(price.call,price.put,pric.call.sd,pric.put.sd))
        }
        cat<-0*(1:N)
                if(WD=='CAT'){
                X.TEMP<-SimTEMP(City,N0,N,Sim,1)
                cat<-colSums(X.TEMP)
                Index<-mean(cat)
                print('Index')
                print(Index)

                price.call<-exp(-r*(N))*mean(CALL(cat,K))
                price.put<-exp(-r*(N))*mean(PUT(cat,K))
                pric.call.sd<-sd(CALL(cat,K))
                pric.put.sd<-sd(PUT(cat,K))
                print(cbind(price.call,price.put,pric.call.sd,pric.put.sd))
        }
        o<-0*(1:N)
        if(WD=='OTC'){
                X.TEMP<-SimTEMP(City,N0,N,Sim,1)
                o<-colSums(OTC(X.TEMP))
                Index<-mean(o)
                print('Index')
                print(Index)

                price.call<-exp(-r*(N))*mean(CALL(o,K))
                price.put<-exp(-r*(N))*mean(PUT(o,K))
                pric.call.sd<-sd(CALL(o,K))
                pric.put.sd<-sd(PUT(o,K))
                print(cbind(price.call,price.put,pric.call.sd,pric.put.sd))
        }
        drop(cbind(Index,price.put,price.call))
}
```

Bibliography

[1] Fred Espen Benth. On arbitrage - free pricing of weather derivatives based on fractional brownian motion. *Applied Mathematical Finance*, 10(4):303–324, 2003.

[2] Fred Espen Benth and Jurate Saltyte-Benth. Stochastic modeling of temperature variations with a view towards weather derivatives. *Applied Mathematical Finance*, 12(1):53–85, 2005.

[3] Fred Espen Benth and Jurate Saltyte-Benth. *The volatility of temperature and pricing of weather derivatives*, volume 7. 2007.

[4] Francesca Biagini, Yaozhong Hu, Bernt Øksendal, and Tusheng Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications*. Springer, 2007.

[5] Francesca Biagini, Bernt ØksendaL, Agneès Sulem, and Naomi Wallner. *An introduction to white noise theory and malliavin calculus for fractional Brownian motion*. Dept. of Math. University of Oslo, 2003.

[6] Tomas Björk and Henrik Hult. A note on wick products and the fractional black-scholes model. *Finance Stochast*, (9):197–209, 2005.

[7] Tom Brijs, Dimitris Karlis, and Geert Wets. Studying the effect o. *Meteorol Atmos Phys*, 88:119–128, 2004.

[8] Dorje C Brody, Joanna Syroka, and Mihail Zervos. Dynamical pricing of weather derivatives. *Quantitative finance*, 2:189–198, 2002.

[9] Jean-François Coeurjolly. Simulation and identification of the fractional brownian motion: a bibliographical and comparative study. 2000.

[10] Don Ellithorpe and Scott Putnam. Weather derivatives and their implications for power markets. *The Journal of risk finance*, pages 19–28, Winter 2000.

[11] Hèlyette Geman. *Insurance and weather derivatives - From exotic options to exotic underlings*. Risk Books, 1999.

[12] Stephen Jewson and Anders Brix. *Weather derivative valuation*. Cambridge University Press, 2005.

[13] A. Kiraly and I. M. Janosi. Detrended fluctuation analysis of daily temperature records: Geographic dependence over australia. *Meteorol Atmos Phys*, 88:119–128, 2004.

[14] Bernt ØksendaL. *Fractional Brownian motion in Finance*. Dept. of Math. University of Oslo, 2004.

[15] Bernt Øksendal. *Stochastic Differential Equations*. Springer, 2007.

[16] Anna Lourie. Stochastic delay equations of ornstein-uhlenbeck type and fractional brownian motion. Master's thesis, 2007.

[17] Paul Vander Marck. Marking to model - or to market? *Environmental finance*, pages 36–37, December 2002.

[18] Sebastian Michalski. Blocks adjustment-reduction of bias and variance of detrended fluctuation analysis using monte carlo simulation. *Physica A*, 387:217–242, 2007.

[19] Yuliya Mishura. *Stochastic Calculus for Fractional Brownian Motion and Related Processes*. Springer, 2008.

[20] Fotini Pallikari and Emil Boller. A rescaled range analysis of random events. *Journal of Scientific Exploration*, 13:25–40, 1999.

[21] Davis R.B. and Harte D.S. Tests for hurst effect. *Biometrika*, 74:95–101, 1987.

[22] Achilleas Zapranis1 and Antonis Alexandridis. *Weather Derivatives Pricing: Modeling the Seasonal Residual Variance of an Ornstein-Uhlenbeck Temperature Process with Neural Networks*. Department of Accounting and Finance University of Macedonia of Economic and Social Sciences.